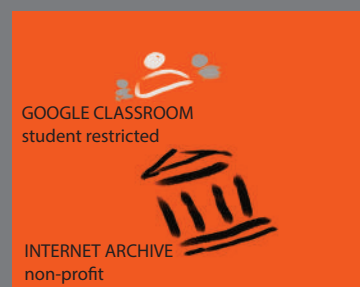
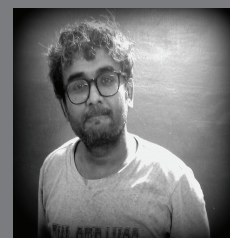


SUBJECT TALKS is a series for university courses of English linguistics at higher learning in a campus environment confused into a better learning for a talk.



Wajdan Raza is a teaching staff at the University of Karachi.



SUBJECT TALKS: talks about language assessment and programme evaluation

How do we connect? SUBJECT TALKS.

talks about language assessment
and programme evaluation

WRITELEFT

۱۷۳



Greetings from Universiti Sains Malaysia, Penang, Malaysia!

I must congratulate you on the depth and insight reflected in this initiative. Subject Talks is not only a creative departure from conventional instruction but also a thoughtful response to the unique needs of learners in under-resourced and socially complex environments.

Subject Talks presents a deeply intellectual and gracefully delivered teaching approach that gently moves students beyond passive reception into a more alert, reflective mode of engagement. Through its narrative style, personal voice, and inclusive language, the method fosters an environment where learners are not merely recipients of information but active participants in a shared, evolving dialogue. In doing so, it invites students to think philosophically, connect personally with content, and experience language and subject matter as living, dynamic entities. Rather than prescribing knowledge, Subject Talks encourages students to arrive at meaning through introspection and interaction, an approach that subtly cultivates higher-order thinking and self-awareness. Rooted in the spirit of Vygotsky's sociocultural theory, this method creates a "zone of proximal development" where learners are gently lifted from the familiar toward more complex intellectual territory. With the lecturer acting as both guide and collaborator, students are scaffolded into deeper understanding, empowering them to engage with the subtleties and nuances of language in a non-formal yet profoundly rigorous way. This approach is especially impactful in educational contexts where rigid structures can stifle student expression or emotional engagement. It challenges learners to move beyond rote comprehension, fostering a sense of intellectual alertness and personal agency. While the poetic abstraction and philosophical depth of Subject Talks are its strengths, they may occasionally veil key instructional insights. This is not a flaw, but an opportunity: the addition of gentle signposts such as thematic summaries or reflective cues could enhance accessibility for all learners without compromising the elegance and richness of the discourse.

In sum, Subject Talks is an intellectually generous, emotionally resonant, and pedagogically sound model of authentic teaching. Bravo for this bold, humanizing vision of education one that invites students to think, feel, and grow. With its unique blend of insight and inspiration, this method equips learners not only with knowledge, but with the critical and creative faculties to engage meaningfully with language and learning.

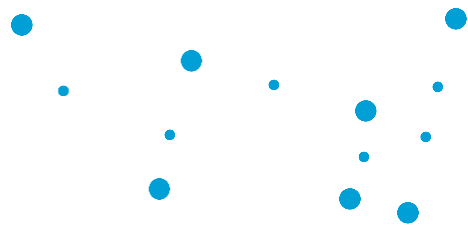
SHAIK ABDUL MALIK MOHAMED ISMAIL
Staff
Universiti Sains Malaysia, Penang, Malaysia

-

● frequency

A band of hair
A girl matches in the city
Lowering deep as the sea
Has just come back again.
A text old, left as a carbon
To corrupt the melody
Finding its ends as the sea
Has blackened the bank again.
It is just a frequency.
The pixelates colour
An eastern lass on a screen
Projecting a reception as an eye
A human wants again.
It is just a frequency.

WAJDAN RAZA
Staff
University of Karachi



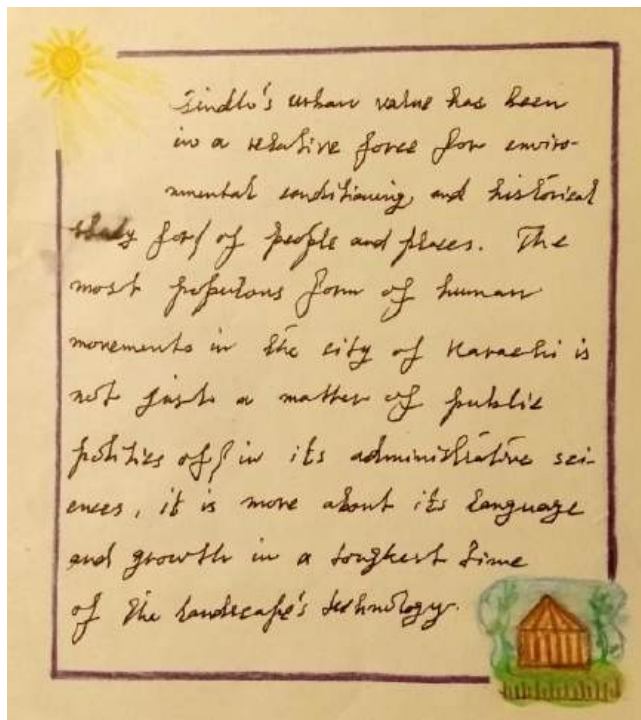
-

WRITELEFT is a creative property.

-

FOR THE MOTHER WHO PRAYS EVERYDAY.





Sindh's urban value has been
in a relative force for environ-
mental conditioning, and historical
study for of people and places. The
most populous form of human
movements in the city of Karachi is
not just a matter of public
policy, but in its administrative sci-
ences, it is more about its language
and growth in a toughest time
of the landscape's technology.

You can read my
handwriting, can't
you?

The colour is
Maryam's. She reads.

I am her uncle.

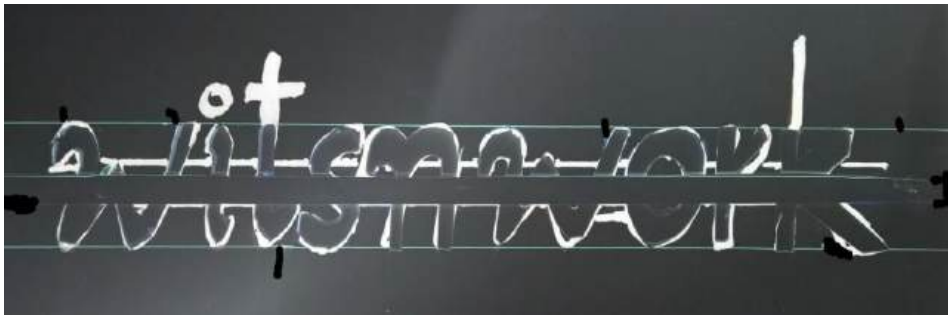


Tooba loves me.

-

RIGHTS RESERVED WITH THE AUTHOR

PHOTOCOPIABLE TEXT



-

SUBJECT TALKS

LANGUAGE ASSESSMENT AND PROGRAMME EVALUATION

Wajdan Raza
Teaching staff
University of Karachi

—

<i>SUBJECT TALKS</i>	4
<i>SHIRLEY AND THE MASTER</i>	6
<i>TEACHING</i>	8
<i>KARACHI</i>	10
CLASSROOM TALK 1	16
CLASSROOM TALK 2	21
CLASSROOM TALK 3	22
CLASSROOM TALK 4	23
CLASSROOM TALK 5	26
CLASSROOM TALK 6	32
CLASSROOM TALK 7	34
CLASSROOM TALK 8A	42
CLASSROOM TALK 8B	43
CLASSROOM TALK 9	49
CLASSROOM TALK 10	57
CLASSROOM TALK 11	64
CLASSROOM TALK 12	74
CLASSROOM TALK 13	79
CLASSROOM TALK 14	85
CLASSROOM TALK 15	93
CLASSROOM TALK 16	99
ASSESSING ENGLISH LANGUAGE TESTING	105
A PAPER TO READ	112
BIBLIOGRAPHY	130
NOTE-TAKING	135

-

An assessment is a measure of quality and difficulty of a language learner as a user in different roles and contexts. The choice of the assessor created to establish and improve a mechanism of continuity of the discourse of learning and teaching and talking and writing seems to have become a formal decision pattern for language analysis and assessment today. The contemporary use and participating process of language assessment are likely to determine a language programme and its evaluation here.

-

SUBJECT TALKS

I was asked to describe my teaching methodology for a visit and a settlement with an Australian university in 2020. The reluctance I had grown in me had required some words for an acceptable justification and its acceptance. The following I wrote on a paper like this one:

Hardly do I deliver lectures in classrooms of subject and language. Interaction has been the key to an initiation of communication with students of Pakistani origin in Karachi of different zones and resources. Recent information transfer to subject learners might have been observed with teacher's and learner's talks in/and beyond specified time scheduled for frequency of interactions in under-resourced rooms. Classroom management is opened with small talks for student-teacher trust and linguistic proficiency measures and built upon theoretical lectures, reading sessions, discussions and academic partnership for a conclusive presentation into student's/learner's rate and a form of transfer of knowledge for the tests and the scores.

A student of mine may enquire about the true value of the interaction in some episodic relationship of information and structural development for a thought and its travel to a subjective mind for a more lapse and its penetration before it is academically received for a process and a differing product centring around it or across its body of communication. The talks he or she listens to are audio clips of my words about the subject and its texts. The language participation's requirement for the students in a teacher's talk is not just democratizing but establishing a healthy relationship in the room where they are likely to sit and enjoy. The language of the books, the journals, the magazines, the papers, the people and the electronic media changes its patterning in the thinking reception of each mind before it is met with the other and many doing the same differently. The thought developing in the course of

—

some subjectivity of life and practice needs its spatial existence in such a room of teaching and learning or just learning as a continuum for both the teacher and the students as true and closest participants of the communicating in the medium of the air. The tradition of fixities of the teacher's place and the students' ones are common and ensured in a local organization of educational practice in Pakistan. The respect the environment enjoys may contradict the conditions for a life communication in public institutions of higher education here.

Subject Talks is a result of an opportunity created for the students studying in very difficult ranges of rooms of teaching in a two-room teaching department of English of the University of Karachi. The room for them is influenced by criminologists studying and supporting the Islamic scriptures for a correction in the city's prisons. English and its space for education is liberal if not contradictory with the behaviour of the people around the girls and a male teacher of some English attitude in it. The humble head of the department of criminology from the province of Sindh had advised me not to teach the girls. This stress of the community, lack of public funds, Islamic space and its constraint for university girls and increasing gender disparity in favour of girls for a support from the WB, the Asian Foundation and many good funding agencies were changing themselves into a medium for the life that they and I wanted to go into.

Subject Talks has a *you know*, a *we know*, an *I think* and a *you* for the talks stopping the monologic linguistic relationship and finding a mental space of togetherness. The pronominal *we* is slipped into it too. They are not the fillers but the subjective relationship of growth and its liberty for the mind that I need to talk to. The audio recorded sessions are received by the students and the transcripts and edits are here.

Subject Talks is not aimed to lecture but change the information into a life learning.

-

SHIRLEY AND THE MASTER

Once upon a time, there was a girl called Shirley. She used to paddle up her pink cycle to get to her master labouring. A grey beard, a long moustache and black, curly head hair had a magical influence upon his bright but assisted eyes lazing in a serious condition of the weather of Sindh. The provincial condition was not just bad because of the mass migrations of the country's families to it but its climatic arrogance upon the speaking inhabitants was difficult to be realized in the parched zone of the master.

"What can we do, master?" she enquired. Shirley was too young to remember the global conditions for political and economic realizations in the living of everyone and everything.

"Can you talk to the reptile?" the master continued. "The snakes have poisons but you feed, you know."

"I do not feed, but I cannot kill them."

"Do you want to kill?"

"No, I do not want to. I want everyone happy."

"American labs have poison centres. They want to learn what poisons the things have for the humankind."

"Master, is everything so precious in the world?"

"Your laziness for the act of the kill is a motivation towards its learning. The being is not just a man. It is the existence. Its absence is its continuity."

"What is it?"

—

"The speech has the man today that the word and the meaning and the meaning and the word are structuring him. You need him."

"Do I need to learn him?"

"Not really. The speech that the being speaks."

Shirley's brown hair was too long to be left carelessly for the eyes. She tied it colourfully and smiled when she planned to go back her home called *The Shirley's*.

Her family members had left for the UK for a permanent residence. She could not leave him. The city's old, clumsily adjusted for her new style of living. A black smart television and a green wardrobe were placed to the two corners of the room open to the dining space reaching the cooking area of some American formulation that everyone could see. The books had been shelved and the tabs had been sent the energy. The communication she wanted was not intrapersonal but connecting, improving and controlling. The master wanted her to see it. She was planning to teach a course and had been given a responsibility in the University of Karachi, a notoriously girls' place for learning the days she was paddling up and down, tying her hair out and smiling for everyone and everything.

Shirley wanted some orderliness before she became a part of the communication that she had developed into *The Shirley's*.

SHIRLEY'S ORDERLINESS

HEALTHY CLEANLINESS	The door to the stole and the bag to the word
POSITIVE SILENCE	The bird to the man and the head to the toes
POSITIVE LEARNING MOTIVATION	The moving dust to the chirping lips
THE SMILES	The eyes to the master
THE OFF-CLASSROOM TALKS	The sound to the speech
TECH GOOGLE RECEPTION	The tip to the sight and the girls to the boys
TECH GOOGLE SUBMISSION	The thinking to the learning
HEALTHY VISITS TO THE CAMPUS	The shoes to the <i>jutti</i> and the word to the meaning

-

TEACHING

Teaching was noble in the religious history because of the traditions of moral values, principled actions and honour in addition to the liberty attached to a human. It has not lost its verbatim but redefined itself in the contesting human population of nearly 8 billion on the planet. A market, supply chain and certified teachers are not the newness of it but a point of realization of information structures structuring the human behaviour for the present and the future. The text written or spoken or saved in some documentation is guided, screened and looped back and forth for the reader or the viewer with or without an idea he or she wants to travel into or sleep walks for. The learner must participate to face the marketization of the relationship developing for their needs, actions and performances.

Is it a matter of a trust between too many groups of humans and things in and across educational institutions? I do not want to know. The truth is the learning is difficult, not because they do not want to come to it but it is connected to the word redefining for a neocolonizing market and some neoliberal controls. The education is a million-dollar work for a small school in the city of Karachi. The English market in or outside the country is far more moneying for an under-privileged or an under-resourced community. Extremely challenged Urdu speaking communities in the city have to have the learning of the schools for a better life in Pakistan if the global citizens want to insist upon the shut-down or stop the *madaaris* (Islamic preachings for the Quran and Sunnah) of Islamic history for a deviant meaning of human life and death and *dunyaa* for the nation-state system. The meaning shift and its problems are too obvious in ethnolinguistic religious speech communities. The Muslims' welfare state must have an insight.

The publications are the past. The new is not seen in the words. The seen is not a sleep walk either. The visual stream of the bright eyes of a Pushtun selling at a cart on the

—

most expensive roads of the city, the red lines in the ones of a Hindu serving ice-cream and yellow jelly in Sindhi-speaking Larkana, and a Punjabi cleaning my house's sideways everyday for his daughter's education are a challenging degree of human literacies for all experiences.

The meaning is marketed. The language is valued before it is processed and replied to. It should not happen in a speech community. A psychic development of a political idea for a control of a mind for its material value may have been a reason for a disturbance of lack of mutual intelligibilities among the language users. Their bodies and brains are an environment that needs a motivation of a chemical change for a growth and liberty.

An excellent person must exhibit excellent conditions for life communication. The ability to teach is teachability. The learning must be in acquisition of the abilities of a true human that can read and reply to the market and the word true to its meaning. The language seen is uttered, worded and expressed for a life for the past, the present and the future. The walks may be synced with the musicality of the language.

The HEC-Pakistan and the Statistics Division have reported numerous achievements of the governments since the country's political liberty. The military for education, Islamic values and technological developments are a significant relative performance of the State of Islamic Republic of Pakistan. Radical left and liberal arts are the placid waters for the business, political and military rich groups. However, the contradictory or inconsistent or decisions by authority for political decisions of political groups and organizations have radicalized the norms and the behaviour of the citizens. The nation must thrive.

-

KARACHI

'Citi Never Sleeps' was tagged on an American card of my brother when the military had established its contacts with the world and the Muttahida Qaumi Movement (MQM), a leading socialist, liberal movement of Urdu-speaking Muhajirs, the latter had replaced the religious numeracy for an unstable democratic defining construct for the city busy in commerce, television commercials and educational toughness for all religious sects and non-religious firmness of life practitioners. General Pervez Musharraf in Pakistan and the self-exiled movement's founding head Altaf Hussain from London were in the international press. An anti-religious firmness was felt to be a part of the city's life to undo Islamic warship against Russia and gain back the support of the historical American people for a moderation of the religious life. The Punjabi and Pushto Talibans were approaching Karachi for a collateral damage. The Muahjirs' unfortunate engagement in the continuation of the Jihad and the city's Mayor's amazing work for structural development are in a public crisis today.

Karachi is a massively large and unimaginably true urbanity for colour, creed, costumes, and cosmopolitan lifestyle against or with an equally increased minarets and mosques for Muslims. Mr Jinnah, the founding head of Pakistan, wanted a modern welfare state for Muslims wherein everyone including non-Muslims could live happily. Mr Hussain's popularity for urban development, equal rights and provincial liberty had attracted the downtrodden and the educated in Sindh, Punjab and KPK. This increase of ethnic liberty to multi-ethnicities of a developing nation for human indexing's improvement in the Muslim world seem to have caused a weak intra-ethnic identity but jumped the movement into a national political fairness of justice for the people. Mr Khan's recent persecution in a state-controlled cell is likely to disturb the ecological frame of the nation. The nation needs Mr Jinnah's democratic values and the state must be seen as a welfare for everyone. The bloodshed's history in the 1970s

—

for the continuity of Karachi as the centre's capital of politics and commerce , the 1980s for the ethnic life, the 1990s for the life's results into a conflict with the religion and its biggest KSA's supported Sunni sect, the 2000s for land-grabbing, the 2010s for lawlessness and the 2020s for social deprivation of the Urdu-speaking communities in the presence of the massive human and money investments of Punjab and KPK has a reason of a multi-ethnic extended family of this man who has studied the world in the University of Karachi, the USM-Malaysia and the KCL-London. The KCL's was a failure in terms of a formal study in the Waterloo Campus, but it has maintained its relationship with me. Dr Nick Andon, a teaching staff, was a good friend of mine on Facebook. Dr Malik of the USM wants to work with me, for national strategies, with or without the language planning, in his odd times. The books I read are American and European. Hardy do I see an Asian Press in my studies for some religious and ethnically different young, white, coloured, politically secret faces of Pakistan waiting for my sessions to continue. My airtime seems to have come to an end. The head I have is a bit damaged. The room the university offers has had a bad voice travel and a weak student-teacher and student-student coordination. I asked a friend of mine in the administration if we could improve the conditions of the rooms. He required massive investments of the governments. A medical doctor of PN Shifa helps. I have had to drive for the service and its payment in Pakistani rupees. The University of Karachi's low visibility for such closeness may have the historical seriousness for or about the community or the provincial head's political naughtiness or the goo-miles-away Islamabad, the shifted centre for the Islamic Republic of Pakistan, or the economic meltdown. The traumatic lives of the university's professors need a serious state's attention. Their residences, medical service and the family support are unrealized in the most spending city. The support staff are unhealthy too. Let me report a recent incidence in one of the unspeakably wrong rooms for a continuous university teaching for a bi-sex classroom population of Muslims.

KU's Bright female student: *Sir, what is the scope of the course of linguistics?*

Me: *Do you want to know about its learning scope?*

KU's Bright female student: *Sir...*

Me: *The life. It you will passes on to your children and family.*

KU's Bright female student: *Sir...*

Me: *The other side of it is its economic liberty for those who work. English has it.*

-

Maybe, she wanted to know about higher education and its continuing part for the one who needs some excellence for a reason. A good national government must differentiate the educated from the poor by offering them a higher living quality. Malaysian government offers a car for mobility, a house for living, medical facility and political support to its citizen receiving higher education. In Pakistan, it is not bad government schooling, it is bad state mechanism. The state language is antidemocratic, non-nation governing when it comes to communication for work at a state-controlled community. The schools, colleges and universities have almost crippled. The power the state and its language asserts could not be sustained by beautifully educated minds as students and teachers. The KU needs a democratic language of the state for the same brilliance it is known for. I have seen many instances here.

A beautiful Baloch for the eldest brother, a sharp Punjab-origin Urdu-speaking Sindhi medical doctor for the tallest one in Adelaide, two white-skinned Indian-origin for two permanent residents of the city and a tanned journalist for an international press for the youngest one are hitched and growing with their kids and relatives. Usman and Tooba, the sweetest gifts, travel to meet and make me happy. It is a city. The two sisters have handsome and hardworking men for their smiling families. Educational structure's informality in the physical setting of the university and political deprivation might have caused a reason to build a life challenge for me.

19th in the Sittal

Once upon a time, there was a man in a paradise called Sittal known for its art, colour and liberty for bright boys and growing girls, smiling men and flaunting women, and flying birds and quick animals. The sound travelled and the sky jumped to the green plant growing a flower all orange to be plucked and picked to her left part of the misty, dirty hair. He saw her in the Sittal for a day of his birth. The sound travelled again rhyming:

Birthday!

It was sent to the eldest bro on his birthday celebrations in Sukkur.

-

SUBJECT TALKS

LANGUAGE ASSESSMENT AND PROGRAMME EVALUATION

-

QUESTION 1. Is the assessment too challenging in Pakistan?

1. Yes, it is because the student as the language learner gets demotivated if they get a bad score.
2. No, it is not as it helps improve learning conditions of the learner.
3. Yes, it is because the assessor knows the learner's role is a bit different from a formal performer of language use.
4. No, it is not as we know that the test-taker's (that is, the assessee) motivation is established with task development and accomplishment.

-

Question 2: Is a young learner different from other speech community members in the description of language proficiency? Think.

-

CLASSROOM TALK 1

Good morning and Assalaamalikum!

Wajdan's here.

The title is Language assessment and Programme Evaluation.

I will also try make an attempt to differentiate the two phrases of the title of the course.

Assessment and evaluation are popular in today's educational platforms. In fact, they are used in non-educational organizations as well. However, I will try to limit myself to the discourse of education containing texts and their productions. We know that this development of the definition and its control over and across the educational institution can influence the decisions of policy makers of non-educational institutions. So, whatsoever is done here or established here can also be followed in organizations of certain values of cultures and improvements.

Why do we really need an assessment?

I think we need some assessment to establish some measure that the teacher—let's take an example—has delivered something substantially good to the recipient of the community of the learners. Let's consider it the first point that the teacher's performance of the delivery is nice, provided that the condition of the communication is excellent. However, we know that this institution that we are into doesn't enjoy that measure. In fact, it fails a lot because of this lack of communication about the environmental condition for the transference of the delivery of the teacher's thought

—

or their practice to the minds and bodies of the learners. So, what is really important? This assessment can also help us identify a problem and make a decision that if something bad happens, it needs a correction in due course of time for some excellence and improvement. It has lots of definition. I want you to go back to some texts.

I'm going to share one of them and see how the same is defined for educational reasons. We know language assessment is really crucial. It isn't that difficult because when it comes to language, we know that it is developed in an environment and a child or a kid or a human being is born into or is pushed into this very environment. He or she has to grow up and think about lots of things and can also express themselves in a chosen or a given language. We, here, can also refer it to a target language. In most educational institutions in and across the country of Pakistan, the language of education happens to be English, but we know that the language changes in certain conditions of minds and places. We can verify. We can assess. We can observe the place. We can observe the person. We can observe the practice. We can observe the culture and this observation is an initiation of an informal way of assessment of the language.

However, the same language that we want to focus upon can help us develop a notion of language assessment. If the kid responds to us as teachers, we can say that the language is being learned. Therefore, we say, 'Have you gotten the idea?'

'Can you produce this word? Can you pronounce this word? What's the meaning of so and so word? And what have I talked about? So, what are those two summary sentences?'

We can say, 'This is language assessment.'

Whatsoever is delivered is responded appropriately or inappropriately. We have to improve upon the condition of the communication and the transference of the delivery of the thoughts developed in certain conditions of the community of the learners and teachers. We know that the community has its own knowledge contribution.

Both teachers and learners become a part of knowledge formation in some conditions of educational frameworks. This article that you will read has a good differentiation between evaluation and assessment. I think this assessment is limited to the

—

performance of the learner. And the term evaluation is more holistic and used for the programme and the overall performance of the learners' community. Please remember the community of the learners refers to the stakeholders. A stakeholder means whosoever is involved in this educational participation. We know that there are teachers, students, administrators and some social family members. For instance, the student's family members, the teachers, relatives, industrialists, educationists, experts and many other beings of interactional life practices are here. So, the assessment or language assessment in particular happens to be empirical value oriented. That is, this assessment needs to be objectified to a truest sense of numerical distribution. It is 1, 2, 3 or 4. On the basis of these numbers, we can describe the performance as the best or the worst or the average or so and so.

We can also have a comparative analysis of the body's and brain's performances on a continuum. So, out of 10, that's the best performance or that's the worst one. And even if this isn't the worst, that's the performance below normative equation. That's the lowest range of the performance of the learners or the test takers and so on. If not the test takers, as I said earlier, that is the classroom itself has its own assessed part. Whatever is spoken is assessed in this context.

The learner's body language, the teachers and the administrators—all of them—can contribute to the assessment of the language. And what about the environment established and continued for a communication between the teachers and the students in the presence of some good administrators? That's referred to as a critical part of assessment in general and language assessment in particular.

I am afraid we can't deviate from this structured participation of the stakeholders in this educational platform. Therefore, it's pure education and academic texts and their production. It can influence the industry or the market. Whatever happens here becomes a reliable example and its duplication in some non-educational and more professional organizations. If we come back to the evaluation in the presence of a programme, we want to see that what it is. We know it is like a computing programme. A computing one has gotten its own logic. We have gotten courses and subjects and subjects have gotten these courses in an academic programme. These courses have gotten their own directions. And they have gotten their objectives negotiated with the subjective performances of human practices in and outside the environment of the classroom.

What is the subjectivity?

The being present in the room of the teacher has their own perspective and may also exhibit their own points of views. We know that everyone is different. Therefore, this difference needs its acceptance and the same difference can be shared with other

—

differences. The differences can have their similarities with the help of a holistic control for social cohesion. This evaluation is more holistic and the assessment is more empirical, but both of them can be shifted back to a positivist's point of evaluation. The positivist's point of evaluation is an attempt of gathering information and actions for a particular specific decision. That is usually established with the help of a numerical orientation, that is, 1, 2, or 3, or 4. And if it is, say, 10, it is a yes. If it is 2, it is a no. This positivist is more quantitative, but the same can also be supported with some qualitative performance. The learner or the teacher can also be evaluated with some interviewing skills. The interviews can be assessed qualitatively. The positivistic empirical participation of communication for a decision for the improvement or continuation or rejection about a programme needs a support in some post-positivist thinking and practice, that is, qualitative, non-numerical way of thinking or subjective contribution for the acceptance and involvement of everyone in today's life of language and communication. Therefore, we can consider this act of post-positivism as a democratic behaviour of a human performance for a decision. If it isn't too bureaucratic or top-down, we will say that something bottom-up or democratic. And that's how a subjective mode of the involvement and acceptance of everyone is realized.

How do we get to know about the two terms: language assessment and programme evaluation?

I think language assessment was used to establish some language superiority in Europe after the Second World War. Some experts were asked to assess their own language. They had to go to different places. They had to check the languages, speakers and their proficiencies. That was happening in Asia in the past. And that happens today too. We have some British and American curricula, but, I think, they have adjusted themselves and accepted some differential ways of communication. Another bad part of it is accountability. No one wants to be accountable. The word accountability can also go with all of these words: language assessment and programme evaluation. Hardly do we use this word 'accountability' for language assessment. It's too hard to accept it for a democratic behaviour. We know that if something is too liberal, it might corrupt the condition of the communication. There must be some way of accountability. If we can move from the word superiority to accountability... If the language was assessed for its superiority after the Second World War, it was this language assessment and its programme evaluation. The programme evaluation or programme evaluators wanted to improve this sense of accountability of the programme.

How could they improve the sense of the 'accountability'?

-

I think they wanted to check the teachers and the students and the administrators. Some of them wanted to focus upon the performances of the students. Some wanted to focus upon the quality of the teachers. We see that this sense of accountability is successful for a realization today if all stakeholders, that is, teachers, students, administrators, social workers and professionals, are involved and checked. For that, they have devised a systemic information collector known as needs analysis or need analysis. This sense of the language superiority for certain groups of people in Europe was the beginning of language assessment and programme evaluation. Probably, the American way was more democratizing later on. American governments wanted everyone to be a part of the system of education. They came up with the idea of No Child Left Behind (NCLB). And this contribution of the attraction towards the system of education was the initiation of the policy for the evaluation. Today, the evaluation is more internal. In the past, when the language was rated for its superiority, it was more external because people or experts or both were asked to come to the communities and check the language's performance among the speakers. It happens in Pakistan too. There are many American or British schools for some non-Pakistani curricula. They come and they check and they rate and they go back. We see that there are many internal evaluation programmes today.

How can we improve upon the terms of the title of the Language Assessment Programme Evaluation?

Dr John Norris has suggested some innovations in language programme evaluation that we can test a theory. If it is clear to us, it needs its policy and practice. And the same can be regularized and rated. And there is also a logic model of positivism. Logic model means that there are some conditions. There are preconditions. There are some set rules. And they must be used for or against certain practices. One way of the evaluation is internally developed participation of human (supported) communication. The other way is to have some technology and big data. The Google Classroom, for instance, can help establish this evaluation of big data and technology. The teacher's talks and the students' performances, containing lots of words and sentences, can be seen, compared, analyzed, and the information collected can be used to make a decision. Let us stop here.

Thank you.

-

CLASSROOM TALK 2

1. Testing procedure has a process for test-takers, testers, administrators, researchers and evaluators. The test items developed have psychometric truncation between easy and difficult, scoring and non-scoring and knowledge forming and knowledge confirming. The screen's difference from a paper test is used to encourage the learner to reel and write and self-assess a quality measure and peer review.
2. A language test is also used as a human fitness test. A kid can differentiate letters, write words and join them to write a sentence. Their visual knowledge is expressed through transmission and transformation.
3. If a kid is asked to write the word for the sum of a five and a seven, he/she is likely to write the following transmissive writing instances:
 - a. twelve
 - b. a twelve
 - c. it's a twelve

Obviously, writing has conventions and human fitness to perform the task. The transformational one is a developing thought about/on a concept, an issue, a term, a word or a problem. It is more challenging than the transmissive one.

3. Listening has a link of comprehension to the sensory motors, auditory lobes and parts connectivities and speech lobes. A human fitness is seen here too.

-

CLASSROOM TALK 3

There is a fair agreement in the discourse that the idea of (a kind of) assessment is in the description of the inseparability of the control of the bias for attention, development and decision in the (language) user's life. Its academic colour, political frame and global recognition are some fascinating points of discoveries. The assessment is hardly proved wrong when the teacher as an agent of the idea attends, observes, interacts and decides about:

1. the learner;
2. the text;
3. the talk.

This ongoing effort of language assessment, say, is recognized as formative, primary, engaging, liberal and democratic. The summative may lack this degree of language interaction if the test taker is left to the items too away from their language and thought. Yan's Chinese influence is a step towards a different Asia.

Please read and study Yan's (2023) test-taker insights for language assessment policies and practices and compare it with the study of Grande, Leishman, and Skilleas(2022) for a Norwegian experience of oral proficiency. The civil-military relationship has been a central developing engagement for language quality measure in specific domains of communication. Its many forms, registers and threads are referred to as ESP today. A fellow of yours from a very department of learning has questioned the issue of money distribution for the national defence and education in Pakistan. I think both must be in agreement of the public welfare. The former is as educational as the latter.

-

CLASSROOM TALK 4

Is it really necessary to have a desk and a chair for a test room today? The answer is pretty unpleasant: the brain tests everyday.

A curriculum is a primary mode of educational display of a group learning of human communities that has turned engaging, connecting and liberating. Its efficiency is measured with language assessment procedures in general. The curriculum's theoretical frame has its influence upon classroom management. The Google Classroom is an example in Pakistan. Its pleasant part must be seen in civil life's testing plans that we want in the nation-state operatives. The procedures ensure the success of the curriculum and the participants. However, its effectiveness is limited to the test-takers as learners and the test-takers as achievers. The impact that it draws secretly has the academic/non-academic and human/nonhuman administration. A machine's contact with a human for a language test, for example iBT of ETS, is fairly intelligent for a large-scale testing today. The relationship between the curriculum and the assessment needs to be strong. Therefore, we need this commitment if we think a syncretic attitude of language is in the possession of language assessment. For example, a kid who may jumble the word's pronunciation /sit/ or as /tis/ or /tiz/ until it—the kid's contact with the world—is rhythmized into the patterns of the world. Another example is a test of English for Specific Purposes (ESP) beyond human communities that ensures a successful radio-communication via sonic (imitation) form of human language for auditory judgements and human and machine decisions in civil aviation and the defence. Is it debatable if some noise is added to the sonic form for a better life in the air, on the ground and under the waters?

Does a human being speak before they listen?

They listen.

What is in it?

—

It has sounds, images and connections.

The brain lobes are in friendship. Obviously, listening is not watching but it does have many broad and narrow human language functions that are different from a machine work and cannot be isolated when the linguist needs the description. A speaking test has an inseparable confirmatory listening, receptive direction. The test administration's requirement for a separability of the reception from the production is rather clinical, diagnostic and tabulating for a language (learning) management decision.

-

Does a human being speak before they listen?

They listen.

What is in it?

It has sounds, images and connections.

-

CLASSROOM TALK 5

This educational paradigm has become rather democratic today. In some sessions of the semester, we tried to appreciate the literary and democratic processes of language use and assessment. I think in one of the sessions and the audio clips, I have tried to comment upon the radio communication that is pretty important these days and even if a mishap occurs in such ways of communication a damage may cause losses of things and humans. Similarly, if a mishap occurs in the production of this language in the particular context of phonological order, miscommunication may occur and damage the reputation of the beings involved in the process of human communication and relationship. This inclusive language behaviour is very important. The behaviour of language assessment is aimed to help the assessor include the teachers and the students. It's rather accepted that this type of the assessment needs an agent and the teacher can play the role, right? Even if he or she can test and teach, we know that this teaching is inseparable from the notion of the testing. When it comes to language assessment and use, I think this inseparability is too serious. Whatsoever is spoken is received and verified and sent back to the brain and the body of the learner for its communication and its success and its improvement for a healthy life practice. He or she is talking about it confidently. This auditory reception and production go together. That's easy.

We know that this role of the assessment, that is, the agentship or this agent is taken by the teacher. He or she can also teach and test together. I think this is what we have agreed upon in some of the sessions of this semester. This exclusive behaviour that I want to share today is about the inclusion of the students. I think this is quite radical because the teacher themselves struggles at times as it comes to the assessment

—

because of its increasing formality for its separation from informal means and settings. It has its value. It has its own decisions that can affect the assessing positively or negatively. Therefore, I think this radicalization of this inclusion of behaviour is a somewhat sensitive, complex and may gain social resistance. It is also referred to as educational paradigm and is something at the centre of the society that can impact the educational practice. This practice can impact the centre. It has a bidirectional realization in human perception of some research and study. I don't buy the word 'society' here. Since it is used in the literature, I think I should also continue using it. I know that this term has gotten its directions like a small model. It is at the speech community' centre. This educational practice with this inclusion of your students and learners can improve this quality measure adding to the responsibility of the assessor. We should ask the question:

How can we do so?

Besides the teacher's trouble in it, the student can also have this serious stress upon their mind. The teacher can help them respond with. He/she doesn't know about this assessment of quality like this that what the teacher wants to say, or he or she wants to have some directions. That's a bit difficult because if I ask any one of you to help me improve this assessment procedure, I think one of us or someone in the same community must believe in some positive relationship of the students and the teachers to and with this process of the assessment. If I believe in this improvement of the assessment with the inclusion of the students/learners/participants of the communication process for and about this subjective question, there is a need of finding a procedure. For instance, one of the students can be considered as a teaching assistant which is quite popular in today's school systems.

The school systems means universities and colleges and I don't know if these primary schools can be included by now. The teaching assistants' popularity in university schools and university colleges has a social role to perform as well. They can be invited to the process of the assessment that we know is about the quality. It is about the measure and this calculation that can predict the future performance at this first point. Secondly, that can estimate what the person can do at present and that's in that given condition. There are many possibilities of this assessment scope. For instance, the learner who is here can take notes in the condition of an informal gathering at a physical setting of a library-like place that has regular visitors. If someone in this group of the learners can take notes and discuss, you can estimate the same can happen again in the same physical setting for the learners. That is called estimation. The other one is called predictability that the person can do this in such conditions in the future. We can predict and express some capability of doing the quality measure. We can estimate that this is happening here can happen again positively and the learner who's

—

communicating here can do so in the same condition at a different point of time, probably, in the morning, like we do it in the University of Karachi, right? Everything is included in this process of the assessment. It has its space and time. You know, these two connections continue influencing the response rate of the learner and the teacher. If the learner is included this way that he or she can share that what *he or she can do, I can do so, I want to do, I can write a story, or I want to write a story*. The learning objectives can be set this way. Let's discuss this with one of the learning objectives. These objectives can also be matched with the quality measure traditionally known as self-assessment. For example, (I repeat here) *I can write a paragraph. I can speak about social values. I can negotiate. I can read some documents of scientific reasons or literary values or argumentative structures in English for 20 minutes. I can discuss it with my friends and siblings. I can stand up and address a gathering.*

I think that's one of the ways of the self-assessment that can be added to the quality measure of the same process as a result of the inclusion of the learners as the teaching assistants. At the start of the semester, I was trying to include the teacher as an agent of the assessment in general and language assessment in particular. Today, we are trying to include the learners as the teaching assistants. These two perspectives can add to the quality measure of the process of the assessment. There is good strength of information gathering for the output, which is referred to as learning outcome. There are two ways. You can also come up with the teaching objective and the learning outcome. The teaching objective can be coupled with this self-assessment sentence structure:

I want to write a story.

I want to read a novel.

I want to speak publicly.

I want to do things in an increased stress.

If the teacher can have a look at all of these self-assessment sentences, he or she can also design their objectives. That's easy. You know that that isn't to do with institutional activity of language teaching and assessment. Please remember this way of doing is not contributing to institutionalization, right? The learner has this way and this desire. This teacher wants to do this. This negotiation can help out of that and achieve certain objectives.

I want to do some IELTS and TOEFL tests.

How can I do so?

What are the ways of doing, studying, taking?

—

I want to write sentences.

I want to correct some of them.

I want to read some paragraphs.

I want to listen to some clips.

I want to watch some English speakers.

I want to improve my intelligibility among those who speak the language differently, for instance, the one from the country you live in.

Someone from Africa and the other one is from the UK may show the difference as well. There are three speakers from different regions that speak the language differently. The language is the same, but the behaviour of it is quite different impacting its intelligibility. This inclusion of different speakers in the same process, obviously, can also influence the learner.

You may realize that the language isn't the same but has changed at different points of the world. I think I should refer to it so that you can connect. I would like to share all of such texts. Let quote from Zahos (2020):

...teachers providing the assessment criteria without involving learners in the construction process which either leads to learners' confusion about the meaning of the criteria or their different understanding of the criteria from the instructors.

It's what the person is trying to say here that if the learner is absent from the process of the assessment, this learning process will be perceived as a negative, destructive, obstructed. If the learner is involved, this whole assessment gets its positivity and the confusion of the learner is reduced to a greater extent. There should be a good theoretical support for the establishment of the learner- teacher assessment criteria. That is, both learners and teachers are involved in the assessment. The former is an agent of it and the latter is supplying their information. This act of theirs can also help generate learner awareness of or about the assessment. The practical justification should lead the practitioners to the concept of co-construction. That is, there is a good cooperation of the construction among the learners and the teachers. The self-assessment sentences of such learners can be used to create a good relationship between the learners and the teachers. This can help the teachers as assessment agents develop assessment descriptors. The descriptors contain descriptions of competencies of the learners as the assessees. Please remember the learner has many roles to perform. One of them of theirs is the assessee. He or she should move from the role of the learners to the one of the assessees. This movement

—

needs the description of the competency, which is referred to as the predictability of the estimation of the performance that the learner will try to acquire. at a particular time and place. Please remember that particular time and that particular place will remain the limitation of the procedure of the assessment. Therefore, most international test units want their learners to perform again and again, not that they have gone away, but the world has changed.

Let's wrap up today's discussion. The inclusion of the learner is central to the co-construction of the radical establishment of language assessment as a quality measure. The co-construction of the radical, let me rephrase this, is a mark of achievement of today's democratic way of language assessment as quality measure. This assessment is a quality measure and that achieves some improved accuracy. Any question? Do you have any question?

Is quality measurement the same thing?

Yes, that is a definitional operation of the assessment. So, if I use it as a term, I want to refer to the phrased term quality measure. That is, it's predictability and estimation.

-

I think that's one of the ways of the self-assessment that can be added to the quality measure of the same process as a result of the inclusion of the learners as the teaching assistants.

-

CLASSROOM TALK 6

Let's summarize the last part of today's session that the three regular participants are the learners, the teachers and the raters. The learners are expected to perform, continue and improve. Their continuation is conditioned with the teachers' feedback. The teachers can give oral and written feedback. If someone as a learner or a performer has written a piece at grade 4, he or she can be given some positive oral feedback. If he or she has made a mistake, they should be given an opportunity of it situating itself in the same linguistic environment for some good fluency and accuracy. If that issue persists, you might help him or her look into and improve.

If you think that this level does not need such accuracy but needs good fluency and sense development in some phrase and sentence structural ways. That's fine. And if this written one can be improved with some oral experiences and practices, you might ask the learners to use their oral practices to look into the (gaps of) writing. Similarly, if you want the learners to improve writing, you can ask them to take some notes, write some words and use them appropriately. That depends upon linguistic and social conditions of such learners. I think this performance can be continued. The learners/performers need positive constructive feedback. And sometimes the same feedback can be replaced with some kind of score like 0 to 10. Someone in score 1 needs improvement; someone in score 9 is an excellent performer. And this performance can be continued in different linguistic conditions. Someone who's done a good job and scored 5 at the scale of 10 can have 7 in his or her next performance.

I think this scoring, which is the second part of the teacher's participation in the process of language assessment, should help guide the students of the course to the last important group of (expert) raters for the language. These language raters are

—

good at finding positive and negative practices. They love evaluating and are also expected to develop some scales or use some scales. These rubrics and skills have certain groups of audience. The learners' performances are central to the use. They look at them and try to improve their performances before they are asked to do it at an occasion formally for some documentation and scoring. If they are provided with such scales, they can objectify their language use. For instance, oral interview's technique may be used to develop and assess the oral introduction. Similarly, writing has its structural growth for information, analysis and expression.

This scale can help the performer improve objectively and place his or her performance at that particular moment of his or her language documentation successfully. The same can be coupled with listening and reading. It may include regions and their language information. The foreign place and its information in a testing language is always challenging. The testing team must help design the test item in a linguistic environment of the testee. You can also witness that the scales for the testee, the tester and the administrator are different.

Please allow me to read some of the sentences from the old Hungerland (1955). The author wants poetry back.

It has long been fashionable to regard poetry as irrelevant to the learning of all languages. It is sometimes reluctantly allowed admittance. It has usually, after the main business of the language learning agenda as being completed as a kind of light entertainment which practical people need not take seriously, but which myself to stimulate interest.

Poetic expressions are complex for a testing unit. The structures are different.

-

CLASSROOM TALK 7

Good afternoon and welcome back!

I think those classroom discussions contain some additional information. You can connect both of them to get to a good text for reading and studying. It is really impressive on the part of student's learning. The text that I have today is to do with one of those topics of the course Language Assessment and Programme Evaluation. We know that language assessment is a very important part of the course. It is dependent upon the testing, the scoring, the evaluating and the grading. To make this happen, we have to administer what the ways of developing tests and what the ways of their writing are. I think this writing part is really important when it comes to the testing in general. It is to do with the language testing. Writing becomes primary. We notice that the language of the testing isn't the language of the learning, is it? The language of learning is certainly different from the language of the testing. Even if we make all efforts through, we do see that this written part—even if it's on the screen, with some movability of gaining pattern—cannot go with the language of learning, the language of classroom, the language of the one is taking it up or the one who is planning to do with it. It's not going to be the same. That is really problematic. I think that's one of the issues of language assessment when it is synced to the procedure of the testing. We know the testing is very formal, right? Even if we think that we can have some tests in a semester gathering, we have to maintain the same level of formality. For instance, if we can't ask the learners to take the final test or the general test before those small tests referred to as low-stakes tests. The other test is commonly known as a high-stakes test because of its finality or conclusiveness or irrevocability. The final is extremely decisive because of the tradition of the testing

—

procedures including the test item development, its accuracy and its liberty for its autonomous operation in terms of linguistic development for the mental spaces of/for the learner as a true test-taker for a performance. The formality for its formal organization can be matched and compared with the learning operations of the communication in the discourse of learning adding to knowledge formation for the participation of the learners, the teachers and the programme contributors. That hits the learner's future makes it the high-stakes.

Let's sum it up that there are at least two basic types of the tests: low-stakes and high-stakes. What I'm trying to make a point here is that one of the issues of the organization of such tests is this inherent separation of the language of the testing. It is hardly the language of learning or use. The language of use is always different from the language of the test. Sometimes, the language tested is good at using the language of the learners, but that doesn't happen every time. It doesn't really happen with the IELTS.

We say: 'No, sir! It doesn't happen'.

Please go to the books of the IELTS. You will have lots of examples. You will notice that the language written there isn't the language of the users even if that encourages them to use his/her language. This motivation of the utilization of personal expression is a bit different from the language written on that piece of a paper.

We know that the language written needs to be standardized, easy, simple, very clear and accomplishing in some time unit scale. However, it lacks this description for the user's language that learns and exhibits itself. The user who uses that language isn't the confinement of time at least if not the space. I think it does depend on both. I'm sorry. It does depend on both time and space when it comes to language testing and language use; however, as soon as it is formed with the structure of language use, we think that this time constraint doesn't influence as much. Similarly, you know, the place that we enjoy inside classroom at a place, say, in the market, is important when it comes to the difference of the places: the testing place and the language use place. This inherent separation of the language of the test and the language of the use becomes a central problem of language test authenticity. The terminological word used for the separation is the authenticity or language authenticity. We can also add this handout that contains some good understanding about the test used to interpret the term the language test authenticity in two ways: the language test and situational authenticity. That the language used must have a good reference to the situation the test taker happens to interact with or through. For instance, when a test taker is assumed to interact with or through a particular language native English speakers in the future, it turns out into one of the points of the list of the needs analysis. This point, I repeat, has this assumption that the test taker is likely to interact with native

—
English speakers in the future. The point must help the tester or the test developer, at least, to have a situation supporting the possibility of the point's occurrence on the day of the test. The language, you know, changes all the way. If someone is going to be employed at an aviation unit, the test developer is likely to have a situation of the same kind of interactiveness. The language will have certain different words, different phrases, different expressions and adaptations. The other part of the term that is language test authenticity—the authenticity—is rather tricky. I think it's a reference to the communication resulted with or from the interaction between the test taker and the test language.

What does it refer to?

What does authenticity refer to?

This authenticity refers to interactional authenticity. The language on a piece of a paper is in some degree of interaction with the test taker who says the moment that interaction begins gets into the interactional authenticity. I think a reading of yours has also described this from the perspective of the test taker's cognition and meta-cognition. There's something that happens in the brain, in the mind. He has this pre-exist kind of mental space. I think that's what he's trying to say. How can we do so? I think that this issue can be resolved when the test, the test developer or the teacher or the one who helps, the testee themselves can have a look at some samples of such written tests or such situations. The difference between the language of the test and the language of the use can be handled successfully. Before we approach this language test authenticity, we can also have a look at the importance of language tests.

Why do we really need them?

I think some of the needs of the test are nationalizing, naturalizing and emigrating. That is, the one who wants to emigrate needs to pass the test of the language in use at the place that he or she is planning to visit for a long time. It's referred to as the need of this immigration. Sometimes, it is also a need of citizenship. The one, who wants to go to Australia needs to pass a General IELTS. Sometimes, this condition is relaxed if the one who is going to the place uses the language fluently. One of the needs of the use of the test is the acquisition of some immigration and citizenship. Sometimes, some universities' administrations demand a good degree of language proficiency. Language assessment, which is quite bigger in scope, happens to be replaced with the term language proficiency or proficiency movement.

The whole concept of language assessment that we know is different. If you go back to the first talks of the semester, you notice that the assessment is not the testing. When it comes to language assessment, it also has something different to offer; but

—

this idea of the proficiency movement has replaced or influenced this bigger idea of language assessment today. It's all because of some political control of some nation-state governments. Sometimes, some governments want one language to grow. For instance, the UK wants one language to grow. We know that the UK has gotten lots of languages today. The one who is there and wants to be there may be asked to take this test successfully. Estonia, a past Russian federation, could not ask its people—or some other people—to take an Estonian language test because of the pressure of the country. The Russian federation wanted the countries to promote a different language and the language spoken in Estonia was less important in terms of political growth. There are many other European countries that follow the CEFR. Hungary has its own test. Austria has its. There are many national governments. They promote their language tests. They hire some testing agencies to control people and encourage them to use their language before they visit their places for short and long time periods.

The language tests are so serious in the world. The same language test has its specific promotions. The language, for instance, English, has its promotions in different domains of communication. The one who is in technology must be capable of using the language in the same domain. The domain must influence the language; the language must not do so. The relationship is quite different. If the domain is passing point of the language, it has its testing. That helps standardize, promote and manage. It is specific speech communication. It helps communities of respective lengths. The test taker's personal information can also help the test developers develop the content. For instance, if the test taker's information contains lots of Pakistani people, the test developer will certainly, you know, engage them through his or her content development because he or she knows that the language of the test is different from the language of the use. Or a similar language use is different from this language test or the language of the test.

There are many studies here. We know that the test has another issue today that if it is too international or if it is too global, it gets expensive in terms of money. For example, those tests that we take here. The ETS tests are quite expensive for developing nations. You have to pay around PKR 50,000- 60,000. It is not a small amount for the pocket of a young Pakistani citizen to be employed for money or ready for business for money. What about O-level and A-level tests? They are quite monetarily expensive. That's another issue of such international tests that they get expensive and the one who wants to go to the place of the language has to sign in before he or she plans to move to the place. The backhand support for such global tests is amazingly exemplifying. It may have made it an expensive affair for us.

Let me summarize. Language tests are quite popular today and really central for

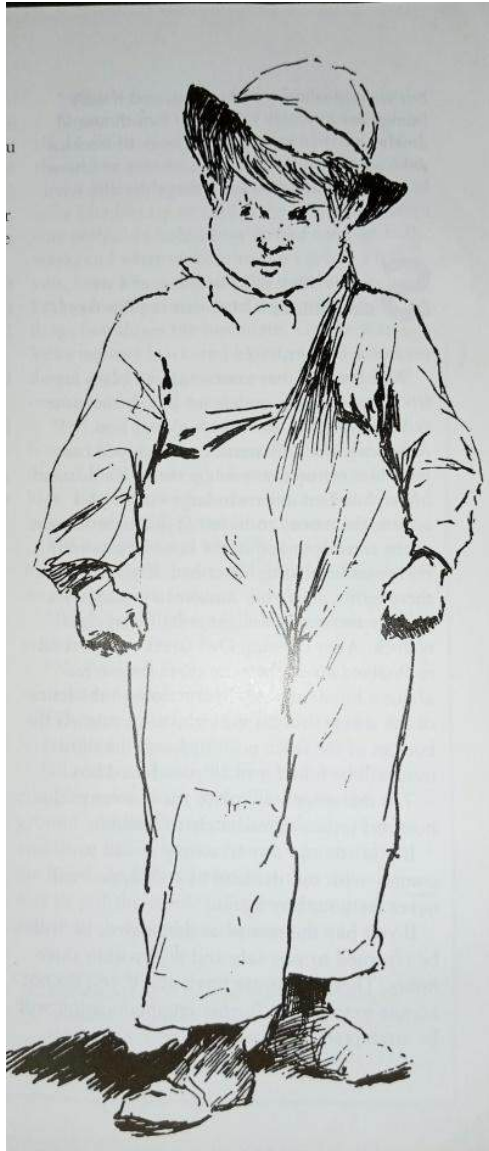
—

global human mass movement. They are used to nationalize and accept other citizens difficult to be matched with language use because the theirs is different from the users. To overcome this problem, such language developers as writers have come up with the idea of the authenticity. That is, it has its situational and interactional placements. The situational is a reference to the domain of the language developed and placed on the day of the test for the test taker. That is, the one who wants to join technology has to take up the test of the same content and value. The interactional authenticity is a reference to this cognitive labour. The test taker as an interactive....

We stop here.

-

Even if we make all efforts through, we do see that this written part of it even if it's on the screen, with some movability of gaining pattern cannot go with the language of learning, the language of classroom, the language of the one is taking it up or the one who is planning to do with it.



-

HANDOUT

Language testing backwash effect is an idea of conviction that it has an effect upon learning and teaching practices among the participants. It is good if the test is pleasant. Its pleasantness is a backhand tedious procedural point mechanism:

1. language approach to language testing;
2. language authenticity;
3. test-takers' information;
4. contextual information;
5. test items' development and accuracy for a predictable measure;
6. testers' team and objectives;
7. language test's administration;
8. testing fairness;
9. scoring and evaluation;
10. scoring and evaluation's interpretation;
11. research;
12. research dissemination;
13. continuation, improvement and/or replacement of the above.

CLASSROOM TASK

Assess a man's or a woman's or a kid's English for the character sketch that the tester has been asked to audition. Use the procedural mechanism.

-

CLASSROOM TALK 8A

We try to make the teachers understand the language assessment agents. The national language of Pakistan has a minority group of Indian origin, as the first users, before it is used in all provinces happily. There are many first users of the language but its political reality is differently placed. Its use is barred in educational institutions in general, not that the language has ethnolinguistic problem, but it lacks the literature of the global development. The confusion about the access to the national language is hard to be removed despite the constitutional privilege its users enjoy today. This talk of mine is in Urdu and has been translated into English by a machine. This is a different, tedious method of the accession. Learning is complex and not limited to the acquisition of knowledge but it is used to help understand and apply. A book *Studying and Learning at University* offers a good piece for students and scholars.

Urdu's cultural values are too seriously poetic and religious. Its use in the religion is fine but it is not used politically that well. Karachi's political meltdown and economic recession for families cannot be neglected for its re-situation in India. Physicists, astrologers, chemists, pharmacists, and many other disciplines of social sciences, arts and the people of architectural sciences cannot use English well.

If you don't verbalize, communicate and make it public, you won't be able to see the democratic behaviour of the language, the traditions, thoughts, and the needs of the people to live in a society. When you don't see all of this and don't see the democratic behaviour of the language and language and people, it will be lost.

Language testing may be used to promote regional languages. The financial and political problems for developing countries, like Pakistan and India, are less severe.

-

CLASSROOM TALK 8B

This hangout has gotten a picture emphasizing many meanings here. I think this character sketch can get your own perception to grow.

I don't know if I can get your perception. I think one of the perceptions of yours is resentment at least, right? That the test taker is not happy with the testing plan that he or she is going to take up a test today and that he/she is going to learn today is not he/she is going to take up a test today. I think these are two different moods of the user of the language. I think one of them is happy in some cases and the other one isn't that much. But the same test can make someone happy. If the testing unit is aimed to promote global engagement and communication, the language of English has been playing its role in helping. If you can have a look at the very first lines written at the side of the handout, language testing backwash effect is an idea of conviction, that it, has an effect upon learning and teaching practices among the participants. The term backwash has its neighboring words to make it more definitive and operational for the course, that is, language testing and its effect. There are three neighbouring words for the idea of conviction of the backwash that there is a backwash that has its effect upon teaching and learning practices among the participants. On the one hand, we know that the learner is under the influence of this impact of the test upon their lives. On the other hand, the same is translating its impact upon the teaching behaviour that contains content development instructions and the interaction with the participants. The test is too formal—probably, the teacher in office is going to use the same language of formalities to enhance the capacity of the learner; and if the language is too informal, probably the same informality can guide him or her. If the language and the test are too structural that contains lots of syntactic arrangements of language expressions, the teacher, you know, can use the same language with such instances of a structural development and

—

its combination. The structural movement is more like foreign language learning taught course of GTM. The test has its impact upon the teachers and the students and the test-takers and the teachers; and sometimes they are also referred to as language participants or classroom participants in today's world. One of us is using the voice and the rest are receiving it. I'm sure that they can also participate positively.

The second sentence is: it is good if the test is pleasant.

We know that it needs to be pleasant. This is what we see in one of our sessions. Please go back to the second handout containing young learners' pictures. They were happy and I was looking at one of the pictures yesterday. An adult was quite happy with the testing procedure in the States and Australia and their happiness was about the shift of manual assessments to AI-based testing involving no human perception on the screen at least. Its pleasantness is a backhand tedious procedural point mechanism.

What is this?

What is this effort-based performance for the achievement of a happy mood of the participants?

The first one is language approach to language testing. The most contemporary one is communicative one which is referred to as communicative approach that the language is spoken and the language in use is the language and the same should be used in all conditions of teaching, learning and testing. This simulation that becomes a sample—a representative part of the language achieved in use of the language in use that we see in the world—is witnessed in the room of teaching, learning and testing. So, the simulation is a reference to a kind of language sampling representing the same at present. We can't go back to the past. It's really difficult but I know that whatsoever is documented is in the past, isn't it? So, it's in the past, sometime in the past or immediate past. We could also consider it near past and sometimes far past. I don't know if human minds want to go with near past or far past; and sometimes they want to go with near past and not far past. Sometimes they want to have the big lobe of the brain involving the past temporal relationship that depends upon the capacity of the intention of an individual human mind. Sometimes, the same mind wants to go ahead with this future part of it that is rather predictable.

That's predictable and that has its predictability. Communicative one is really, really fascinating and may be used in this one. There are many approaches. That is, a structural one that the language has as a structure and no language can be spoken beyond the structural part of it. If something is communicative, it has its structure and the same structure can also be used for its functionality. On the one hand, we have communicative approach; on the other hand, we have structural one. To make this

—

structural work, there should be something functionalism developing or dependent. Linguistic functionalism is discussed in some Australian universities that this structural part is really important to someone aware of this functionality of the language. Something that is communicative and beyond the scope of this awareness of structural part of the language is equally important when we want to deal with the structural way of it in the presence of functional concepts of language use. What is this functional concept that we use that we have a cell phone with us. It has its lots of software packages. We have buttons or icons at fingertips with us and try to use them. We try to use our brain to use a site; and we know that the moment we use our tips to scroll or tab or just do something with the screen, something will pop up and move. The properties of the device can be used to help the brain of that individual user's function. Cell phone is in operation, we know. It has gotten its own properties and characteristics, the software packages and some other syntax. The same syntax of the software can help the syntax of the human function. So, the syntactic—one the structural way of language—is in operation in the presence of functional concept of language if someone is in contact with such a cell phone that contains its own syntax. The functional element is a meeting reason between language use and its communicative requirement for a reality as a means and as end (Martinet, 2006). This is the first part, that is, language approach to language testing.

We have to make a decision about the approach of language learning use, teaching and testing. I think that the language authenticity influences your choice of the decision. Is it interactional or situational? Is it both? We know that the situational is easy in. Someone who is doing technology should be asked about the same. The domain should influence the language, not the language the domain. We know that the language has such immense power. Therefore, we will know that this permission of the domain encourages the control of the language upon each and every body of the land; therefore, it isn't the authority of the domain. It is in the authority. It is the allowance that the language control needs. That's the second point. It helps you pick up a choice for this decision.

The third point is the test taker's information. I talked about this in Urdu some minutes ago. The test taker who's going to be your guest should provide his/her personal and public information about age, gender, studies, languages and (expected) visits to foreign lands. He/she will use this language and this test performance and he or she is in contact with these people in Europe and we've made a contact with other people in other groups in the future. Is he or she available for research improvement and the test item development? I can, under this third point, add to the second point of language authenticity and the fourth in connection with this contextual information. We need the test taker's information and want to share and gather contextual information. That is, where is it going to be held? Where do we want the tests to

—

happen to occur or recur? Who will be the participants? Who will be the administrators? Who will be the people around the test takers? Language testing content development is under the influence of some degree of language authenticity that is advanced with improved with test taker's information and contextual information. Let's have these four points together to get to the fifth point. What's happening around the test taker? The test developer must be aware of this contextual information that the one who is taking up the test is in contact with. The people and things and he or she will be in contact with them and things in the future. Then, this is the set of these four points should help continue the discussion about the rest including the fifth one: test item development and accuracy for a predictable measure. These four points, you know, should help us move to this backhand work of the test development that should have some test items. Sometimes, some contain one item, just a question. A question means that the test taker is asked a question to reply to and that it is an invitation to an extended written response comprised of a number of sentences and one or many paragraphs. This writing is remembered as paragraph writing or essay writing or article writing in a traditional learning and teaching practice. We remember this as a subjective response. We remember these where we as language teachers and testers do not buy it all, right; we do not believe it; but we know that whatsoever is available to the observer has its objective value, not subjective one. What is spread is concentrated upon one of the sets of the ideas of the expert if something contains lots of sentences and paragraphs and looks subjective though is objectively available to the mind of that expert. We as language teachers and linguists would certainly accept this concept of subjectivity. The point's first part is test item development we may show our agreement with; there are many other test items; for instance, a multiple-choice question is placed or the test taker is asked to write many sentences or/and just fill in, with the help of some choices. That doesn't require writing. Writing is the translation of the brain using the upper limbs. It isn't just cognitive; it has its own physical operation. It isn't just intellectual participation of the writer; it is a manifestation of his or her physical fitness. The accuracy for a predictable measure has its consequential relationship and some sort of continuum of measurability that some performers will perform at the last sight of the continuum having a zero or a one. Some of us can fall at the mid of the same. A few of them can travel to the right of the success. There will be a concentration of the performances as scores or the spec of the numbers or the scores over the same. We must have this predictable part of the range of the performance of all these learners or users as test takers; and for such an imaginative contribution of such a test, we need an excellent team of testers. We will have to think about the members.

Can we have our members from the department of physics? I don't know. Can we have some from the department of chemistry? Can we have some from the department of psychology?

—

We'll say, 'Yes'. We can have someone from the department of psychology. We should have many from the department of human studies. We should have a couple from the department of education. We must have some representation from pure sciences, research and arts. The arts means that they don't write; they don't speak.

The seventh point is language test administration. Let's introduce it in the last minute of the discussion today. This administration has time, place and people. If such a test should be administered in 30 minutes, 60 minutes, 200 minutes or should be sent to the test taker in a condition of almost time free that doesn't require any time, just go back home or just sit here and have a look at the test and perform. There isn't any time constraint upon the stress of the test and the test-taker. The language doesn't want to go with time here; it needs its place; it needs its person; but most tests are administered in certain frames of time because psychologists love it. The world is timed naturally. The language is timed for a reason too. The learning tension linguists don't believe in today, do they? Don't believe in this learning tension. You know there's just the problem of the psychological tension of language production for a coherence and its sustainable reception and its growth for communication. We believe in the psychological tension, not the time for language and its continuous amazing performances. The psychology or its follower wants the test takers to learn every moment and, you know, slip into a (traumatic) behaviour of life at times or, sometimes, you know, some excitement, gratification, happiness and glee. We stop here. Thank you.

-

We need the test taker's information and want to share and gather contextual information. That is, where is it going to be held? Where do we want the tests to happen to occur or recur? Who will be the participants? Who will be the administrators? Who will be the people around the test takers? Language testing content development is under the influence of some degree of language authenticity that is advanced with improved with test taker's contextual information.

-

CLASSROOM TALK 9

Let's discuss language testing. Please remember that language teachers as assessment agents can't help informally this much. We try to use the idea of language testing. I think as language teachers and classroom managers, we must make an effort to include these teachers as language assessment agents. As I have said earlier these teachers are always available in different situations in the context of English language education. We know that these teachers have gotten their and the language's aims. Sometimes, these two complement each other. Sometimes, they oppose. Some teachers have gotten some national perspectives to the use of the language. I think the perspective can also be grounded on some conceptual development of the spread of the language. Let's make a reference to the course of World Englishes and Multilingual Education in Postcolonial Contexts. Since you have studied these two, you can also use the understanding of the courses for the support of this choice of such teachers in terms of their theoretical development for the use of language and its promotion among participants. If these teachers as language assessment agents can't help that much, we try to use the idea, this concept of language testing. The testing has its own huge scope to discuss in the course of Language Assessment and Programme Evaluation. It has improved a lot. We have seen the commercial introduction of AI.

Let me share a personal experience of the AI use for language and its interpretation. I worded the following happy words poetically for the eldest brother living in Sukkur:

-

19th in the Sittal

Once upon a time, there was a man in a paradise called Sittal known for its art, color and liberty for bright boys and growing girls, smiling men and flaunting women, and flying birds and quick animals.

The sound traveled and the sky jumped to the green plant growing a flower all orange to be plucked and picked to her left part of the misty, dirty hair. He saw her in the Sittal for a day of his birth. The sound traveled again rhyming:

Birthday!

The youngest one used AI for its description in a reporting form:

This appears to be a poetic and imaginative passage that describes a vibrant and idyllic place called Sittal. The story follows a man who encounters a woman with striking orange flowers in her hair on his birthday. The passage employs lyrical language, metaphors, and vivid imagery to create a dreamlike atmosphere.

Here's a breakdown of the passage:

- Sittal is depicted as a paradise where art, color, and freedom thrive.
- The man encounters the woman on his birthday, and the sound of the word "birthday" is emphasized.
- The woman's hair is described as misty and dirty, with orange flowers growing from it, creating a striking image.
- The passage employs rhyme and repetition, with the sound traveling and the sky jumping, adding to the poetic and musical quality of the text.

Overall, this passage seems to be a creative and imaginative piece that explores themes of beauty, wonder, and celebration.

The punctuation for parallelism is American English influenced. The market of numeracy has the language and its tailored performance for the users. The stylistic variation has been introduced for more than a natural language performance. The moment the idea was introduced to the market, the same was just grabbed into the same field of learning and researching, that is, language testing. This testing has its own market as same as we see it in educational industry. The seventh point that we left with, at the end of the last interaction in the classroom of linguistics, is language test administration. This administration is more academic in nature, that is:

—

How much time do we need to give to each and every test item? Do we really need to give the time bar to the completion of the item?

Sometimes, we have to ask this question:

Do we really need to give the time bar to the test item and the user or the test taker?

I also said in the last interaction that this time frame tension is in the issue of the linguists is more to do with some psychological framework. However, we know that the psychological framework is in connection with the language use. It's very difficult to separate the language from the psychological development, obviously. Do you support me? If you go back to your course of language acquisition and learning, you know that the psychological development of the user is very important. If one wants to assess this language proficiency, that's true. Do we come to this artificial environment that we refer to as a kind of simulation's intention of the use of the language sometimes, I think the psychological grip over the whole component of the test is a bit unfair. Because of the same, as soon as it comes to adult education—it's not young education—this one-to-two-hour activity is not that significant for language testing and proficiency. The language assessors and testing services units have to begin offering some relaxations.

One of the relaxations is the time bar that the test taker can be offered with some kind of flexibility. For instance, the one who is doing or taking it up on the screens should be given such opportunities of time utilization that he or she really wants to at home or in his or her office. The test-taker who wants to finish a test item on the screen and does it successfully can leave the screen for some minutes and do the work as convenient and come back and continue. Isn't it happening with the Google e-forms these days? I know that this e-form activity is not synonymous with this procedural performance of language testing, but it doesn't negate a good backwash effect of such language testing procedures. The time is one of the administrative decisions that the tester really wants to have, before he or she develops the item.

What is it that some items should be aimed to receive some extended written responses of their test takers? More than 10-12 sentences in a go, and the same bunch of sentences can be spread over different ideas and their placements in a kind of form of paragraph forms. These are two administrative decisions that I've just introduced you to, that is, time and test item development in terms of subjectivity and objectivity for a good result of the (objective) assessment. There are many other administrative decisions; for instance, the test is pencil-paper based or if the test is screen based—if the test is rather linear or non-linear... The linear means that the easiest one will come at the start of the test and the most difficult one should be placed at the bottom of the same set of the test tasks. This is frequently observed in such test material display.

—
Sometimes, you know, the test items can also be used as the test tasks. Is the test going to promote the communicative performance of the user and his or her voice about the world, the person, the people?

You know, the item's subjectivity has a changing language life to the community and its objective pattern wants validation, confirmation and reconfirmation to the language proficiency. What is usually received is seen objective for the reader or the tester or the administer as he or she needs to process the performance for a product of a rank or a number that must have the elasticity of generalization of the language and its production for the proficiencies and their acceptance in and across the speech communities.

What about the text size of the test item? Is this clear to the test taker? Do you allow the test takers to enlarge the text for reading? For instance, if you can do so, they should make a request to the administration that we can't read the small letters and if you can have some bigger ones. You can come up with exclusive font sizes in addition to the regular visual reception of the reading texts.

I think that's again academic administration linked up with some administrative administration. We have gotten at least two administrations here: academic administration and administrative administration. The administrative one includes the funds, electricity supply, land resources, some technological equipped places and simple places containing chairs, tablets, sheets, pencils, pens, erasers, rulers and some reading texts that if the test taker can't read the test items in the absence of some reading text support. The administration should be able to support such beings with some texts. That's also referred to as the assistance. Do you really want to relate it with Google Assistant? That assistance is not that linguistic; it is more psychological. Obviously, you know, the reading text given to the test taker should not be copied over and through with the pen or the pencil of this test taker. If you can sum up the same, this test administration has gotten two types of administration: the academic and the administrative. The academic administration has some academic decisions. The test item is going to be subjective in terms of the reception of extended written responses of such test takers. And the other one is about objectivity, containing some questions and tasks having no written responses. That means that these responses might contain some choices and words and phrases and some fill-ins. The first half of these administrative decisions and academic administrative decisions is a kind of the other one. It's exclusive nature of the test taker's request.

The third is the medium of communication streamed by some pen and paper-based, right? The fourth is the use of the medium for language testing improvement. That is, if the test is AI-based, can we have this psychological, psychometric involvement in testing performances? That is, you know, AI has this psychometric performance. We

—

do have it here too. The easiest one may come at the beginning and the most difficult one may come at the bottom of the test items, because the one who has gotten some energy can save it and use the same instruction as to the performance of the last one placed and appreciated as the most difficult one. I think that's the easiest way of doing. Sometimes, some test takers, if not good, can place the most difficult at the beginning and easiest one at the end. Or they have their own concepts; it does depend sometimes. They have gotten their own theoretical and practical justification and may want to pick up the most difficult at the beginning and the easiest one at the end.

What's the right order? That isn't about those right orders. That is about this tradition that the participants have to follow. For instance, you know, here in Pakistan, it's really difficult to come up with direct communication. If someone asks a question, the respondent delays the answer. He or she really wants to talk about something else. And the one who has asked this question may feel some uneasiness or some negative feelings at times. It is a face in sociolinguistics. A very political form of the violence of the term and its connotative meaning is exploited for some legislation for women rights in Punjab in 2011-2014. Sharmeen Chinoy's *Saving Face* has a meaning of the culture and its people. The face in linguistics must be differentiated theoretically though. That's the tradition that the speaker has to follow. These testing performances have theirs. Someone who is going to take up the test should be aware of such traditions.

This eighth one is fairness, one of the most crucial parts of the testing for the improvement of the process of the transparency. For instance, the test item should be protected and coded and should display offers to the test taker on the day of the test. We can claim this. It's a matter of the transparency. Let's have it secret and out of bound. I think all international agencies follow this secrecy of the test items' descriptions as a way to this transparency that may contribute to the fairness. I think the other one is that the test must not be negatively biased. We know that it always is in it that we can't escape. Everyone has some good degree of the bias. Sometimes the bias can go negative because of some reasons. The two of them are personal and public. The personal means that the one doesn't want to appreciate the other one and is remembered as the personal bias, a negative one. If he or she does it, it turns out a positive one. The public one that this normative behaviour of your community can help the test taker go to or oppose the other.

That nothing changes. It's all the same. If you really want to do something different and wonderful, a bit different from the content of the experience of the test taker, it should not be accepted. A risk or a twist is my way for this test development. Well, the risk isn't available in the language. The test writer may continue:

'Have you ever seen the sky? Does it contain the stars and the moon?'

—
'I haven't seen that.'

For instance, that's why that isn't accurate, is it?

Even the beginning is all impossible or at risk. It is realized that an 'at risk' is an ill-combination of words and ideas. All of these sentences cannot be used for academic justifications in academic writing. To make it clearer, the language must be simple and contextually rich for the test taker or testee. If someone wants to do it the other way, the non-plain or an 'at risk', the testing fairness should allow the user to do it for a procedural requirement for its qualification for a regular use and repeat for the sake of language improvement, proficiency and individual liberty; but this normative behaviour of the community can resist.

That is, the individual is in opposition with the global wisdom about the language. There are many other steps to the fairness. The tester's bias, the administrator's bias and the test taker's bias can be normalized. I think all of these groups of language testings will have their respective biases and the researcher who wants to analyze and improve the condition can normalize these biases for the sake of the language and the testing fairness. This fairness can also lend itself to a pre-administrative measure. For instance, the test taker should be involved in the testing measure.

'These are the test items. Please do so at home.'

This pre-administrative is in practice in the IELTS. That is, you make a request for the test and they accept it. They ask you to make a payment and send you some books and some reading texts. You prepare and may ask questions. They try to reply and help and have the tutorials. You go to the test centre. You take the test. And you wait. You get their reply. They can also give you some samples. And even if you want them to show you a paper, I think, you can make a request; but I don't know if they have this policy, but they can share some samples with you. If you have a score this way, you can have this much score.

You perform to listen. It is both a performance and listening. It's like this. Writing, listening and speaking have some skills to monitor across the language in question. And what about reading? Reading too!

It's like some speakers, in the IELTS tests, are used for this production of speech. The test taker needs to receive this spoken information for some micro processing, in some conditions of facilitation and turbulence. The turbulence is the public reception and physical conditions resulting into a lack of international coherence of some standards. The individual one that we see in the ETS Internet-based testings is an excellent degree for a human fitness test.

—

And then we have the post-administrative part. That the one who has gotten the score should be in contact with the same agency that helps him or her to it successfully. This isn't just testing; it's a product. It has its own company interested in its users.

It wants them for various reasons. Maybe, it needs their write-ups and human relationship. I don't know if I should use the word 'company' here. I think the user and the company are permissible. I think the British Council is like a company, isn't it? It has its affiliation with the government of Great Britain.

The extraordinary ninth point is the scoring and evaluation. Let's just get the last point for five minutes. I think the word score has the value of kind of performance. It isn't a number. It has a score. It's a number, you know. If someone has scored 7 out of 9, this score has its performance value reflecting the actions of the head preparation for the careful movements of the upper limbs for writing, the eyes for the sight and the mouth and the nose and the ears for the speech and the ears (and the eyes) for the hearing and the listening. This is a formal behaviour of the language user for a language test and language testing. I think the terminological dependence of the linguistic score as high as the value of the performance has the consistent repetition of the response rate developed with the test-taker, the rater and test and the administration and global recognition. If someone has scored 7 out of 9, it is a utility in the created isolated conditions of communicative performance from the environment. The theoretical specification of the semiotics of the numerical value is less significant than the issue of the linguistic evaluation of the language and its users in the specific conditions. That is, the 7 in the IELTS band system, an indication of his or her achievement, should help him or her get some laurels, some universities' invitations, some organizational recognitions, some social mobility opportunities and personal growth. That is how a number becomes a score to reason. The academic reasons indicate global utility of English use or language use. Thank you.

-

Do you allow the test takers to enlarge the text for reading? For instance, if you can do so, they should make a request to the administration that we can't read the small letters and if you can have some bigger ones. You can come up with exclusive font sizes in addition to the regular visual reception of the reading texts.

-

CLASSROOM TALK 10

The handout is about language testing's backwash effect. I repeat it is an idea of conviction that it has an effect upon learning and teaching practices among the participants. It is good if the test is pleasant. Its pleasantness is a backhand tedious procedural point mechanism—it is a thirteen-point moderately controlled here. We have discussed its eight points. We stopped at its ninth's beginning, that is, the extraordinary scoring. The point has the phrase the scoring and evaluation. We know that the scoring is a kind of a numerical performance in some situations. It is a kind of a number that depicts some physical performance. I think language teachers and language approaches' practitioners will appreciate the term because most approaches' practitioners and teachers are likely to motivate their users as the learners to use the same activity of a performance.

To perform is to speak and write. They want their recipients to be productive that degree that whatsoever is in their mind and the brain should be expressed. That is, I think, one of the biggest challenges of the day because we know that there are many types of learners and one of them is the one who doesn't want to speak. Maybe, he or she wants to read a lot. He or she may also miss it too, but we know that such receptions at times can't be done formally. The one who is here can observe the environment and process the lobes of the brain. The reception is really informal. Similarly, the one who is at home can read different texts and relate all of them to a common objective of a certain course in a certain language. What I'm trying to say is that this type of learners isn't going to perform that much. It has its extraordinary situation, but communicative (approach's) teachers and those who want to follow such approaches of production sides of the language—that means that they want their learners to be exhausted on the side of the production of the language in the States—really want their learners, their teachers, their experts, their researchers and their

—

intellectuals to exhaust themselves for a language tech and its overarching control and governance in and through a certain language.

It's also happening in today's educational environment. You know, we want our learners to perform a lot. We want them to be rhythmically balanced at the beginning of their language acquisition. If they can do so, with the help of their head, then the neck, the primary mode of balancing or pairing up and some other parts of their body. If the body, you know, gets into some rhythm, the same can perform linguistically. It's all performance from the beginning of such learners' educational, contextual, practical world of knowledge. Write and speak and read and listen and see and touch and smell. All types of reception, you know, should be translated into certain degrees of linguistic exhaustion that we think can retain some unique characteristics of language creativity.

What is language creativity?

Language creativity is defined with the help of infinite size of utterance productions. The one who is creative and born with this creative part of life needs this linguistic exhaustion. It is depicted in educational institutions and common in university environments that we want productions from the sides of the learners, the students, the scholars and the intellectuals. I think, we want the same from our learners as language users. Therefore, the first part of the ninth number of this procedural mechanism of positive backwash effect is fascinating that it contains the scoring. It gives some performance. It gives some competition. It can also incentivize the philosophy or philosophizing competition in the presence of another attractive opportunity that may be referred to a cooperation that the one who is competitive can work with someone else. It's fascinating. The same competitive environment can help generate cooperation. We can refer this to learning cooperation or cooperative learning. This cooperative learning must be seen in the context of competition upon language production with its exhaustive linguistic value in almost all educational institutions of the day. Only the same number can retain such performing values and units. For example, one of the learners capable of walking up to the place of his or her teacher can have a point that you know he or she has understood and motivated themselves and walked up to the place of the teacher and really wants to address the gathering for a minute or two. Even if he or she can't do so, he or she should be given some points and units of performance because it is referred to as preplanned linguistic instruction or a representation of language creativity that the one who has come to the place has to be blessed with and the one who walks up to the place and addresses with some disorganization should have some better units of performances. What we are trying to do here is we try to control this linguistic behaviour that we see and receive with the possible plan of the producer that the one who, you know, makes this command comes to the place and addresses the gathering even if he or she is

—

assumed to be in a certain organization. This organization is in sync that community's way—perceptual and experiential. You can have lots of points if the one who comes here and doesn't talk should have some points. The one who comes here and talks in an awkward way should have some better points. The better is in the sense of its fitness, not its gradual likeliness of an increase of numerical improvement. And the one who comes and addresses and talks clearly should have some more points. The one who comes here and addresses and talks unclearly but he or she has some good points in some developmental stages should have more points. And the one who comes and talks and shows some confidence should have some points too, but his or her content should be realized. The point to the audience is very remarkable but improved personal, philosophical, exemplified or duplicated. There are many ways to the description of the first part of the scoring. The second point which is the last point of the ninth point is the complex evaluation.

The evaluation is to do with some positive backwash effect. The effect has its influence upon learning, teaching and administrative practices. Learning practices should be aimed to improve for the skills of the language in general. Teaching ones are those that include such skills' orientations and opportunities offered to the learners; and administrative ones help improve teaching and learning practices. For instance, if the one who knows that if he or she does this much can have these points of this performance. Another instance, if the one can read can have, say, 10 points out of 100. He or she knows that if he or she can reach two pages. And if he or she can answer one of the questions set on a screen of theirs can have 10 points too. So, 10 plus 10 will have a 20. The learner's standing is a somewhat different from the process of the evaluation. That's just one of the ways of getting to its description. That is, the self-assessment that we see here can contribute to the description of the learner's evaluation. For instance, that the teacher who knows that this learner has performed this much is for us, to say, 80 out of 100, can have this language proficiency. The learner can also be interviewed. The score 80 out of 100 can help evaluate for a decision for a particular time and place. This quantitative description of or about the scoring can be coupled with a qualitative part of this assessment process. On the one hand, we see the learner as the user, as the performer who can measure their performance on the same continuum of 0 to 100. On the other hand, we can also encourage the teacher, the tester, the instructor and the evaluator to couple the qualitative support with the score of the learner. This quantitative attribute of the learner should be qualified. This quantification that we have been stressing upon should have a point of pass or exit. This point of pass or exit is a certain human behaviour that can't be realized unless and until the one who sees this whole process as a qualitative—biased. Please remember qualitative bias is a positive. There is always good positivity of (the) bias.

—
Where do we get this positivity of (the) bias from?

We know that this positive bias is in the experience in global environments... in the psychological improvements...in relationship with this administration of language testing. And then another part of it is unseen administrative. It academically does help us and or the learners. The first point is their individual performance measure with the learners. The second point is the teacher-tester's qualitative section of positive bias with its point of exit of class or band or threshold or standard, and the third is administration that has academic and administrative allowing the execution referring to the meaning of the execution of the learner's or the performer's scoring measure and the teacher's or the tester's qualitative point of pass or exit.

I think the ninth leading to the tenth is discussed to an extent that the scoring and evaluation's interpretation should have these points of the learners, learner's performance, learner's scoring measure, learner's self-assessment, teacher's or tester's qualitative paradigm, teacher's or tester's qualitative point of exit or pass, teacher's positive bias drawn upon a pass, a pass test and experiences. The administration should allow this to happen. I think that this is a trichotomic kind of a relationship of the administration, the testers and and/or the teachers and the language performers having some research value.

What is the research?

We know that the research needs some reading and some good questions and problematization. The problematization is organized to refer to something observable that there must be a benefiting problem in it. It isn't that just social problem commonly used in daily life communication in local communities of the men and women. I mean that there must be an agreement of a way to theorizing a relationship. It might be a linguistic relationship. For instance:

She talks.

There is a good linguistic relationship.

You know, *She talks.* is two-word in the first place. There is always a good relationship when it comes to the research. There is a relationship for researching. There is a good relationship of the researcher with his or her past practice and a good relationship of the past with the present. And there should be a good relationship of the present with the future. That is how the human brain can process mentally. I think all of these ten points may be used to place the research. It has its description of effectiveness in research books as well. The positivistic limitation may be highlighted to a better degree of social performance of a score or a fail or a pass.

—

There are at least two types of the testing: norm referenced and criterion referenced. The former is a reference to social acceptability of the tester's performance and the performer—the one who is speaking, reading, writing and listening, should be absorbed into the immediate community—probably another community that the tester wants them to be in the future. I think that the other one, the latter one, is of the criterion referenced, has its academic value. That the performer must meet up the requirements of the testing approach. Both of them, I think, can complement and reject each other. I think they should complement. Similarly, the testing has its research environment. No test or testing as a concept can be realized without its research support. The main test item will be the test. The test items are very known, easy and somewhat difficult. The easy items can be separated from the difficult ones. Sometimes some items are less accurate and some quite accurate. And those that are quite accurate can be repeated and recycled. And those that are less accurate can be improved and reintroduced and measured. If they can help the test taker perform better, we can have good responses of theirs. Its content, its question and task and its other ways should have some representations in learning materials.

Is the test aimed to measure four regular skills of English language? Is the test aimed to test all of these four skills of English?

The materials should have good representations of the same skills' tests. If any one of them is missed or weakened, this positive backwash effect will have its conceptual contraction that its effect that is so big, as big as this room is metaphorically, is likely to be reduced to a man's or a woman's. If the material that the teacher, the tester, the instructor, the user and the learner interact with or through that doesn't have all of those four skills' learning procedures for the testing measures, the same idea, that is the positive backwash, will have his conceptual reduction of the size of their room of liberty and control and improvement will be contracted to the size of his or her or their strictures.

Samira is going to be an IELTS instructor. I don't know when she's planning to do so. Even if there is an IELTS instructor and he or she knows that the materials in the room or on the screen used effectively will have this negative impact of the testing. In other words, this positive backwash effect will have reduced conceptual influence upon the environment. This reduction is felt to the individual of the same environment. The individual can resist or show resentment. You know, what's happening here. For instance, a situation is set to describe a condition of a learning problem that everybody could have a test after 20 days and no one has improved listening and speaking. The problem is the lack of doing on the part of the learner as the user for a performance for a measure on a specific day at a specific place for a specific time for a specific content generation. If it results into this trouble of someone's liberty, the

—

positive backwash effect will have this conceptual reduction again. And the man who has seen this reduction can do some repairs in their personal capacity. And if it doesn't happen here, he should do it himself. If he/she can do it themselves, the same positive backwash effect will have its value. It is strategically positive for a testing output.

-

There are at least two types of the testing: norm referenced and criterion referenced. The former is a reference to social acceptability of the tester's performance and the performer—the one who is speaking, reading and writing, listening, should be absorbed into the immediate community—probably another community that the tester wants them to be in the future. I think that the other one, the latter one, is of the criterion referenced, has its academic value.

-

CLASSROOM TALK 11

Good morning and Assalamualaikum!

Positive backwash effect is in your recent memory. Please use it. The very last part of it contains a task to finish. If you can do it in pair, it helps. Let's have a look at one of the publications placed at M M Photocopier. The other one is in your possession at the moment. This one that is in your possession will be discussed as well.

The one that I have is with the photocopier is about some introductory comments concerning English and its testing procedures in the neighbouring country India. We know India has lots of similarities with yours even if it has some historical political differences with the nations. For instance, the country that I've just mentioned has Hindi and Urdu for the context of your land for a national agency for communication. This linguistic condition isn't that different from the one that Pakistan has been into and through. When its school system is almost patterned on some practices, you spend time here. You're doing it at this moment. For instance, this formal education has 10 plus 2 years in institutions of learning in here and at the same patterning of formalization of learning has been reported in the publication that I am trying to make a reference to.

The language, you know, has its historical richness in India. I think it's as same as we see here and its richness is a bit higher and the community living is multilingual. English has its four regular skills because the language spoken is used with, but its testing is a bit different even in the states of India. For instance, its utility as a library language is observed and problematized. It is documented and written for documentation and read. Most educational practices are pushed back to reading courses and subject-based performances, but we know that the language has its own performance in all domains of communication. It's really difficult to accept that the language part of communication or language in its totality can be squeezed into the concept of the subject. Here, we have resisted the concept of the subject used for

—

different courses of different disciplines because the language has its performance in all subjects. For instance, the subject of phonology has a language to use and perform, doesn't it? And the same, I think, is practised in your social gathering. If somebody wants to be engaged in one of those gatherings, he or she will have to use a particular type of the same language. This is really impossible to separate this component of human performance as linguistic from this communicative paradigm. The language cannot be squeezed into the concept of the subject as performances in all domains of all subjects even if it is your English phonology, a non-subject part of human communication.

What is happening in India isn't different from Pakistan. For instance, the language used is treated as a disappointingly library language. Here, the characteristics of such a language are amazingly as far as the documentation's significance is rated. I was reading Latin's language teaching book lately; it had many words of today's technological developed notions of vision and distance for a virtual reality. It may be used for publication and shipment to stores and traditional reading libraries. This library language has two components for its ready performance: reading and writing. The learners and the students are asked to do the same. While doing the same, they are tested the same skill orientation. The listening and speaking in the publication reports are relatively new to different systems of education in India. Both of them are internally assessed, not centrally organized.

What is this internal assessment? What is this external part of it?

Let's take the example of the University of Karachi.

It has an internal system of the assessment. It allows its teachers doing their work on campus, who have their learners tested on different days of the semester. That is, the on-campus teachers are allowed to have their students as learners tested on different days of the semester learning.

What is it?

That means that the university encourages a reduction of the high-stakes tests to some low(er) ones. There are two types of the tests here: high-stakes and low-stakes. The high-stakes are summative in most conditions that they just summarize the performance of the learner and the formative in the low(er) ones helps generate the feedback, quizzes, assignments and projects. And some other ways are used to develop some learning feedback. The same feedback can help improve their performance size. Therefore, such testing procedures are both low-stakes and formative. Writing in general is more formative. The students are asked to write some sentences and show the same to their teacher for feedback and their expressions' improvement. They come back and show the same and prove their developing writing

—

paragraphs and some pieces of writing. What is written as a sentence can be developed into a full-length expression containing hundreds and thousands of words. I think this is what you were doing a week ago. The formative testing, as a result, allows a good reduction of the high-stakes' problem. For instance, do they do something wrong in the learner's performance, they may experience a negative socio-academic relationship. If the tester has done something wrong on the day of the test or in one of the parts of the pre-testing procedures, he/she has to see this issue on the test day that he/she can lose their relationship with the participants involved including the administrators. You have done something wrong here and this test is—say 100 or 1000 units. Doesn't it happen in the TOEFL?

It happens with the high-stakes American TOEFL and British IELTS. We know that this is good because you just enter the place, you perform, you get the score and show your happiness to the world. On the other hand, if something bad happens on the same day and everything gets worse for your near future years—there are many reasons—if you aren't feeling well, say, and if you aren't feeling like performing on the day of your test, you aren't able to write any test experience of negativity in the recent past. I think that the same can impact your expression. You can't focus upon your test items. There are many reasons. The tests you are referring to because of their high-stakes nature have been criticized a lot, but they are well managed and centralized.

If I can summarize, there are two types of the tests: summative and formative. The summative ones are placed and offered at the completion of some formal learning situations and performances. The formative ones are about your regular performance size and reception of the feedback for the improvement of physical performances, written assignments, reading comprehensions, talks, listening, podcasts and discussions. They can have some score and the feedback; and then sometimes, you know, some performances are just sent back to their recipients and the producers as recipients with no grade and with some grade sometimes. Sometimes they don't want these numbers and just want the feedback. I think machines have also been doing this performance if you log into the BBC. They have some quizzical performances and offer you some reading comprehensions and some video clips to watch. You can go through and see if you can fill in the blanks, reply to some close-ended questions. If you can do so, they give you the feedback in no time.

You get your feedback and see that it's much accurate in some situations at the same weblink. I think that kind of performance is the formative rather than the summative here. The summative is more formal and placed to the end once he or she has finished the formal practice of a part of the language in a different domain of communication, like 100-mark test after five months or a year with no feedback or the feedback with no written feedback. I think the one who is doing it receives the feedback in the form

—

of some numbers, some gratification that this much they have earned, say, 60 numbers out of 100 or 80 numbers, 600 out of 1000. It's all summative and I think this is also the high stakes. The high stakes aren't good for those communities that are weak. I know the community that we interact every day is weak. It's weak in terms of political strength, economic liberty and psychological motivation. I mean there are many reasons because the community, you know, has to experience lots of difficulties in this developing nation of Pakistan. Such high stakes are easy to administer because of language testing centralization and support. There are many types. The formative tests are rather decentralized, but the teachers are asked to have some physical performances, written assignments, projects, listening, comprehensions, reading comprehension, podcasts, discussions, theatrical performances and musical ones like song and lots of other things that can generate some momentum of this learning community within and beyond their world of classroom practices. Everything that I've just addressed here should be kept into the concept of the formative assessment. We know that this liberty is a bit different. It's a bit away from this conceptual orientation of centralization of language use—the language spoken in some parts of the English world has regularly been referred to. We know that the centre is really important. That practice can improve mutually intelligibly. This way of decentralization is used to promote individual liberty, democratic ways, personalization and speech and motivation towards human engagement for and about knowledge and its dissemination. A multicentric value should be added to governmental agencies of education, shouldn't it be? The other option to grow into a neoliberal free market structure is here too.

Can you give me your quick response? How many tests do we have?

I repeat there are two basic types of the tests: high stakes and low stakes. Both of them have their effectiveness, but the low stake tests are frequently used for the operation of individual liberty and human voices for a learning environment within and beyond world classroom practices; There are two types of the tests again: formative and summative. The formative test helps enjoy the liberty. The summative one is used to reject it, with or without a teacher or with a course. He or she wants a learning projection. This learning projection needs this assurance of some accuracy that the one who has done so has understood what the course developer wants to say or do or wants him or her to do to this directionality that is within the concept of the course. The course is aimed to set a directionality of knowledge wealth and its richness back into the lives of the participants.

When India is doing wonderfully, it's moving from two skill participation to the three and the four to the listening and the speaking. Then the language that acquires library language becomes a living one that we use for interaction. Its management is easy

—
because the internal assessment which is rather formative can help the teacher and the learner generate some good momentum of learning within the community at and beyond the place. India, I think, is fairly successful because it has introduced a trichotomic relationship: the language of the local, the language of the state and the global language of English use for human communication.

India and Indian people use the language for the language of instruction and communication. The language of instruction and communication, I think, are quite synonymous in their learning procedures.

I do see some serious difference between the language of instruction and the language of communication. I think the language that I'm using is a bit instructional here, but at the same time, it is as common as we see it in the language of communication. It's a bit different upon the test and its day. The language becomes well instructional rather than communicative in its domain expansion. And even if that retains the sanctity of language instructions, it has its richness of communication. We have to prioritize if it is instructional or communicative. We know that whatsoever is instructional is also communicative; but what is not can also fall back to the first one, the former one, that instructional one. The one can also be a bit non-communicative communication that can have this instructional component of language use.

Both TOEFL and IELTS (Academic) are high stakes proficiency test of English language offered to those individuals as users of the language who want to perform and get a score for some placements in universities and professional institutions. TOEFL iBT, I think, is more psychologically developed part of language use while IELTS' language is more academic text generating. Both are popular these days and can be used synonymously for certain conditions of language use. The easy and challenging test items continue changing in computational arrangements in the iBT. Since they test on a specific day and time, that significance increases this variability of the high stakes suggesting a meaning that the one who develops the test and the one who is administering it and the one who is taking it up, under the influence of the sensitivity of time and place and the content that the test needs to generate on the same day.

Sometimes some test takers aren't good. They feel some anxiety and can't concentrate on some of the test items. They can read but they can't listen to; or they can listen but they can't write. There are many abnormalities on the day of the test when someone wants to determine the quality of the test taker and these four regular skills. This variability can also be noticed; I think, it has been reported. These administrators of the respective test have also proposed some enquiries from the side of the test takers who can also make a request for a reset or a retake. This policy, I think, is taking its shape elsewhere and everywhere. I think the tests are quite

—

expensive in India and Pakistan even if these test items' developments and testers and test developers are from the same countries' centres and peripheries, but this effective part of administration controlled from the centres of the inner circle needs its own denomination.

That is a bit tricky. As a result, language testing can also be influenced with or by this monetary part of language promotion.

Thank you.

-

I repeat there are two basic types of the tests: high stakes and low stakes. Both of them have their effectiveness, but the low stake tests are frequently used for the operation of individual liberty and human voices for a learning environment within and beyond world classroom practices. There are two types of the tests again: formative and summative. The formative test helps enjoy the liberty. The summative one is used to reject it, with or without a teacher or with a course.

-

THE TALK CONTINUES NEXT WEEK.

This isn't that simple. We know this because we have similar multilingual interactions in our own provinces and regions. It isn't just Sindh, is it? We know that there are many other languages spoken here. The administrative purpose of the provincial autonomy is flawed upon a single use of Sindhi for nation integrity and its control and power on or over the people. What we are doing is we are trying to follow the same relationship for modernity that has become questionable because of this recent and swift human group migrations and settlements. The province of Sindh has many languages today; for instance, Sindhi, Urdu, Pushto, Punjabi and those who speak Sindhi in the province of Sindh believe differently for language labelling. There is a good potential of linguistic complexity for the decision about the language used within the territories of such Indian states and Pakistani provinces.

Hema (2008) has used Tamil language as an example of the trichotomic relationship of language use. Their work language is used for classroom instructions. British historical exogenous expression of human behaviour of or for or with English is introduced to the same groups of learners in formal settings of K-12. The birth language, I think, has this liberty of settlement and re-settlement; however, the native one has a concentration upon a place with its regular context with respective people. This language used at one end of this Tamil regularly used in one of those states of India and the other one, the third one, Hindi or its national integrating one seem to have sandwiched the second language of English in good developing communities of India. The functional English is good.

Here, the writer views that this English is sandwiched between the birth language and a national language. It receives influence from both ends. And, to observe such a situation that if the test is one skill based, two, three or all four regular skills based; one has to see the picture of K-12 educational system of India. The country has three types of educational catchment. The first catchment consists of public schools or

—

government schools because they have direct support from the government that has its own policy to execute through its administration. The other ones popularly known as private schools are referred to as aided schools that take assistance from the government. For instance, Durbin in Karachi's Karimabad is an aided educational place because it looks like a private one but it needs assistance from the University of Karachi and the Government of Sindh. It's an aided school; it isn't a private one and neither is it the governmental. It has its own liberty but needs assistance. It has its interdependence. The first one is rather dependent; the other one is interdependent and the third one is independent; that is, private school(ing). The educational catchment knowledge and its richness infused into the bodies of the participants are extensively studied in Asia.

Hema's (2008:386) views:

about the knowledge base required for conducting assessment in education have been influenced by socio-cultural approaches towards teaching and learning and the growing recognition of the social role of assessment in education that no testing procedure can be realized optimally in the absence of the variables of culture and society and the review of educational learning theory from behaviorist to cognitive social constructivist models. Shepherd (2000) observed the gap between psychometric testing environments particularly those employing high stakes standardized tests.

Most tests that we do in classroom are non-standardized because they don't follow the procedures in general. I think that's a controversial again; but we know that standardized tests have the reports every year that these tests with these items have come up with such accuracy of responses and, therefore, they are referred to as standardized. There's a good movement's documentation measurement improvement of the high stakes' standardization because they can't afford such mistakes. Therefore, they need such standardization. If a test is centralized but non-standardized, it needs this standardization; but what's happening in classroom can go beyond this concept of the standardization. It has good collaboration with student-teacher and students and, in particular, those high stakes standardized tests like testing cultures.

This means the high stakes standardized testing or tests are based upon testing cultures. TOEFL is based upon a certain testing culture and, therefore, is referred to as a standardized test in those countries where the language is spoken as first birth or native. Contemporary learning environments follow Vygotsky. Vygotsky theory is based upon constructivism that there is this learning mechanism initiated and theorized from a quantifiable learning achievement of difficulties from a zero to one,

—

one to two and two to three and three to four, so on. This constructivist behaviour is the development of the language. It doesn't happen every day because the kid that is developing this language psychology can come up with amazing expressions that he or she can't retain in the future. I think we follow the same Vygotskian constructivist behaviourist's cognitive socio-based relationship of and about language use for knowledge growth in the K-12 or K-14 or K-16 or K-18 or K-22 or 20.

There are two types of the tests: centralized and decentralized. The standardization is the same procedure that we were discussing in those last learning sessions with the handout. If you follow the same in a centralized one, we adhere ourselves to the mechanism of language testing's standardization for a positive backwash effect. There mustn't be uneasiness for the tester or teacher if the action is conducive to learning.

Here learning is perceived as a culturally situated and mediated activity with a language occupying a fundamental role as a mediating tool. The learning community jointly co-constructs knowledge and teacher's feedback plays a central role in supporting and promoting a student's language learning (ibid.).

Student language learning is different from language in general. There is a language with and without human interaction and there is another one with human interaction; that is, a student's language is in the language that is there, but it can have some similarity with them. The one who is assisting and mediating can push the same through the language that exists. We're here materially. Materially means that the language in the book is the language in your talk, the language in your theatre, language in your constitution, language in your commerce and trade. These are frozen forms of communication that can be used as an opportunity of physicality or material performance of human growth through language, through a way.

-

CLASSROOM TALK 12

I would like to talk about some material evidence of English language test items' development and placement to the test takers. You know that the test aimed at the course can be divided into two parts: subjective and objective. I know that the definition and its differentiation are difficult in terms of their conceptualization.

This subjectivity is about a perception growing in a process of confirmation. The objectivity is a kind of simplification of some linguistic occurrences. Something that's perceived differently when produced similarly in terms of linguistic behaviour gets objective. For instance, if you're asked to produce the word *speaker*, most of us will reproduce the same. So, some point of understanding is that the production of the word is the same that the machine can identify and match it with its available corpus. Here, I think, this production of this linguistic combination of sounds as for names seems to be quite objective in organization because this match and recognition over the machine is simpler and the corpus is defined because of its enrichment in a particular array of the system. Here, something that occurs in records and this perceptual development is likely to be (condensed) upon a point of similarity. This point of some condensed similarity can be referred to as objectivity. What I'm trying to say here that this difference of the test's division is really difficult for linguistic knowledge and language studies. We know that the test knowledge is not limited to linguists and language practitioners; it's used in different disciplines all over the world. Its bipolarity helps adjust the elastic movement of the brain and improve semantic control over the idea through a language. It's frequently used with some psychological interpretations of the knowledge. This knowledge's addition to the language interpretation is amazing. Each and every part of the production even if it looks similar is likely to have some uniqueness back in the mind and the brain of the user. Nothing

—

is objective in terms of linguistic knowledge. We know that even if its behaviour is too similar because of its configuration into the corpus, it enriches a particular memory of a machine or a being or a community. So, this is all about this creative part of language that needs its recognition. So, the first part is subjective. We know it's easy as a result that we can produce unique combinations of words, phrases, sentences and paragraphs for and about topicalization of communication as a response to linguistic stimulus.

Sometimes, some test takers are asked to write a paragraph or an essay and comprehend an issue and address the replies to it. Sometimes, an issue that's rather problematic needs its resolution into the paragraph form of expressions of the test taker. That is really subjective. You know, it's too subjective because it's gotten lots of words and phrases and sentences and paragraphs and unique ways of placing vocabulary items on a piece of paper or this finger-typed text on the screen. Because of so much newness, a text even if it is recognized is there—if a machine can be used to match it with some writing patterns, can't it be?

We know that when you go back to objective test items, they are easy because they contain some sentences with some gaps to fill to match. And they are used to invite the test takers to go with some options as best choices. They can opt for any one of them as the most appropriate, if not so correct option or choice to reply to the stem's left or right. Sometimes, some gaps are created within. Then it looks like a sentence, right? And sometimes the same is left with a gap that needs to be filled with one of the best options offered by the test taker's creative response. It is not machine mediated. And sometimes, whether the same stem—or the sentence in this case of the gradual growth of the idea overlapping the bipolarity—has no gap and needs a response from the options developed. So, if any one of them that the test taker feels is the most appropriate is opted as a correct one.

If the same opted one is found correct, it matches with the corpus and the language back in the memory of the machine or the human brain. A human being has their memory that can be retrieved to match or reject the information supplied to the test. That's easy, right? If I can summarize, we have gotten two types of test items for the test-takers' responses: subjective and objective. And the difference between them is really tricky, complex when it is inevitably connected with language studies and linguistic knowledge. I want you to comprehend the same difficulty from the perspective of the central configuration of analogical awareness of the choice of language to the pragmatic function of the same. The subjective one, you know, is likely to have long responses. These responses are usually referred to as extended written responses.

—

So, to get such responses, you know, the tester needs to develop such a test item all engaging, stimulating. To do so, the tester needs to offer an issue or problem or situation at times. And then by doing so, the test item gets its place in a good situation of the mind and the place. Somewhere I read that even if you're going to have some extended written responses as subjective ones, the testers should have some sample responses to frame other testers' mental capacity for the right scoring. The best response representative of such extended written response may be used. The same best response may be shaped into a benchmark as a way to supporting with such written responses.

If the best response can be matched—not matched here—or compared with the other written responses, this best one in comparison can have, kind of, say, 5 out of 9 on one of the responses as scripts, or 3 out of 9 or 7 out of 9, say. And even the best response that is the best response that you think is 7 out of 9 is bypassed. Such a response can attain, say, a 9 out of 9. So, this activity of the tester is referred to as a pre-testing procedure.

An issue is created in the score situating itself into the content. Such a test is usually sent to some pre-university, university or pre-advancement class. Paragraph writing is really challenging because it needs some quality production. Now, let's go back to the test's second part: objective item development for the test-taker. The same objective one needs pre-testing or pre-administration. If the test item is really accurate, it gets you the same response at different places. The sameness is more orthographic for a machine and semantically controlled for a human user. This linguistic production is more negotiating and engaging than the match. Public placement of linguistic abilities has this issue to be confronted successfully. I know that the objective testing is controversial because of the words used. J B Heaton (1979) introduces it in *Writing English language Tests*. A problem of such a test of objectivity can be understood with guesswork. You know, some test takers may haphazardly guess and get a score. And if they do so, the authenticity or validity, if not the reliability, is challenged and labelled as dubious. However, these test items of objective nature are popular among psychologists and psychometric experts. They are also popular in English language test today because they can be developed and pre-tested and found accurate. They can help the testers make decisions. If they can make do so, it's a good way of the assessment. The decision, you know, needs representation of actuality. If the representation is not correlated with this language proficiency, this whole mechanism is likely to fail. Lastly, this objective test item needs a contextual relationship. I know some test items are too decontextualized. I think there are some very popular test centres using socially decontextualized MCQs. They want to use the same again and again. Heaton writes that this decontextualization of language can't help assess communicative part of the test takers. Thank you.

-

We know that when you go back to objective test items, they are easy because they contain some sentences with some gaps to fill to match. And they are used to invite the test takers to go with some options as best choices. They can opt for any one of them as the most appropriate, if not so correct option or choice to reply to the stem's left or right. Sometimes, some gaps are created within.

-

POST-TALK COMMENT

Please visit the link containing an e-form about MCQs. The same has gotten the two weblinks: one from the University of Waterloo for MCQ design and the other Skimpflation for the language at the BBC. Visit them. I think one of them is about a market product and its performance among consumers. It has lots of information. I think the same that you need to watch has 3 minute-plus information. Now, if you're going to ask your test takers to watch a 3 minute-plus clip, you need to see if they can really pay attention to the same from the beginning to the end. Sometimes, you know, the tester wants to show 30 seconds out of the 3 minute-plus watch of a listening. And they continue doing so. The test taker is supposed to watch 30 seconds and go back to the test item, attempt it appropriately and go back to the same watch. This continuity can help generate right responses and their management. The second part of it is that you just help them get ready for this listening part by supplying information to the setting of the content. Later, you just ask the test takers to watch the same.

You can repeat it in classroom testing.

I want you to see if you can show a 3 minute-plus watch of listening clips to your test takers in one of the best settings of English language testing. And then you should get some examples as precedence for such unique behaviour like this. I don't know if the subtitles are really used to test assisted listening in adult learning testing. We do use it in language acquisition in formal educational settings. Listening is usually separated from this visual impact in general. I think it's more administrative than academic or daily life one. The one who has gotten all such senses can use all of them to go to the listening part or this auditory reception. Sometimes, the reception doesn't need. The subtitles or verbatims need this confirmation in the literature of language testing. Thanks.

UNIVERSITY OF WATERLOO and BBC are at BIBLIOGRAPHY.

-

CLASSROOM TALK 13

There are plenty of instances in today's English language test. The user has to perform to be accurate in grammatical composition of the language. To do so, the tester has to plan what kind of test items needs his/her recognition in such test situations. Heaton's chapter 4 in the book is used to introduce many ways to the establishment of the importance of grammar. It begins with MCQs or multiple-choice items. We know that the grammaticality can also play an important role in the set of choices. Although this act of the tester is more decontextualized in language orientation, its predictability in terms of the precedence can help the tester guess the use of the language in context in the future.

There are many other ways of doing so. I think one in the same chapter is error recognition items. Sometimes, some sentences are generated and placed to the user as a test taker. He or she is asked to identify the error—I think this error recognition is a reference to his or her linguistic knowledge about a specific expression. The most common ways of error recognition items contain derivational, inflectional, orthographic and syntactic replacements. Sometimes, some utterances and sentences can be jumbled up and generated and placed to the user as a test taker for an order. He or she is asked to identify the error. I think this error recognition is a reference to his or her English knowledge about a specific expression. This error is not the fossilization of the language user; it is rather a limitation of the language corpus and its permutations extended to the human user for a creative stretch of the expression of or about or through or a language. A red highlighted part that we see in a document match in MS-Word establishes as an error.

Sense completion is a much finer way of language use in a grammatical context. Sometimes some words are placed at the beginning of an idea as a sense developing concept. The same needs to be filled with appropriate words. Even the gap to be filled

—

for a sense completion is less objective in terms of its calculation or match word with the user, the writer or the computer or the machine, as the creative contribution of the language has its control over the human body if the tester has intended to persuade the test-taker to exploit the liberty of an individual expression. The transformational behaviour of the language is determined with a stylistic knowledge, say, active-passive combinations and direct and indirect gravitation to a language use and its differentiation of the subject shift. Sometimes, test items of the sentences need completions with replacements, substitutions, additions and/or subtractions for a (broken) language.

The grammaticality is broadly fixed into semantic and syntactic information. The syntactic is a sheer fixation of a human behaviour for a community contact. For example, the phonemic activity of the speech production is in a fixation for a phonological control generating meaning. It cannot change its place. However, the semantic activity is complex and changing. The phonic configuration into the meaning imagination is a slot developed into the lobes of the brain of the user for a community contact and a communication health. The environmental contribution is likely to have more descriptors of it. The grammatical test item is easy for accuracy achievement in general. However, the objectivity is a challenge in terms of its specifications to the communicative use of the language.

Alderson (2007) has written that most tests are hard to be interpreted in terms of objective specifications. For instance, the language used in language test can also be problematic. Sometimes the test item placed in the language can't be approached appropriately by the test taker because of the language and the context it's wrapped into. And this issue becomes more problematic when the language testing or the test is witnessed. The moment it is used to address child acquisition of language and its testing, it acquires its own challenges.

What are the test development and the challenges?

I think I should repeat that one of the challenges is the language itself. The language used in language test item development can be problematic and distract the developer from the specification of the objective. Secondly, the context where the language has been wrapped into is unlikely to be guessed for the improvement of communication between the tester and the test taker. There are very good language learning websites displaying a content using the technique of the match, ordering, colouring, filling and repeating for a novice. The tests they show are shaped into the learning mode. That is excellent, but the formal behaviour of the community contact is too challenging and complex. The assessment has the best and the problematic fits for the test design at a centre of language testing. This is what is seen in the ETS and Cambridge Testing. The web content is more learning and exciting.

—

When we go back to the test or the test items of a grammar, it is pretty clear that there are hundreds and thousands of test items developed in English language education and the same can be used. I think the most popular ones are MCQs and then there are some error recognition items. I don't know if the CEFR, a popular Common European Frame of Reference for Languages, can be used for the diagnosis of strengths and weaknesses.

There are plenty of examples of language opportunities in classroom learning and testing situations. And sometimes those who score better in a particular skill cannot be generalized for other skills. Sometimes, a spoken skill can be generalized in some context, say, workplace context or the one who is good at the language and one-to-one discussion can also write successfully. And the one who can't do it can have some problems in materializing his or her thoughts into words on a piece of paper. So, the language that the learner develops has its complex coding. That is, that the one who is born and grows receives the language in its complexity that way that has its own visual hearing or auditory and some environmental contextualization with this growing phase of life that is getting closer to the one who has to spend some months and years.

Alderson (2007) writes:

In summary, a body of evidence is developing that shows that the dimensions contained in the CEFR itself do not describe language development. Given that the CEFR is based upon a view of language learner that sees them as social agents, given that perspective taken on language ability in the CEFR is sociolinguistic and its emphasis on text type, discourse type, task features and the like, and given that the CEFR is deliberately intended to be neutral as to target language (there are no scales that describe how specific aspects of individual target language might develop that) then perhaps it isn't surprising that the usefulness of the CEFR for diagnosing language strengths and weaknesses is rather limited.

Sometimes, language tests are difficult to be treated as diagnostic. Diagnostic one suggests that it's really difficult to predict the test item developed if it can be used and has its limitations for a placement in language difficulty. The same test item can be used for a placement test or a proficiency one. There isn't any problem and its utility for two different types of language tests. We know that a placement test is used for an achievement. Sometimes, they can also be used for entry and exit situations of learning. And the same can be used for the achievement. Sometimes, your test items are the same but the tests are different. For instance, the placement is different from the achievement. And the diagnostic is different from some other tests. Eastwood's Oxford Practice Grammar has a set of test items for a diagnostic help for English grammar. Probably, a repetition of an idea in different expressions has a good

—

confirmatory decision for a problem or a pass of the response learned or produced by the user as a test taker. I think there are at least six to seven types of such tests that we need to recognize. Dimova and Yan (2022) prefer local to localized tests in their editorial of SAGE journal lately. They are concerned with the local problems' solutions with both separationists that think the assessment values' global recognition and addition to each act of the test item's formulation and their controversial liberty of idea and its penetration into the procedures of the testing in a certain context of time, space and peoplehood and integrationists who see the quality production of communicative patterns with the involvement of each and every individual's experience for language testing.

Yet, one thing we need to focus upon that the test items must have task orientations. The CEFR and most European based language proficiency models of languages are aimed to improve language user's conditions in their sense of completion of tasks. A test item can also have lots of tasks. The task needs to be taken up and finished and that gives them some sense of achievement, completion, personalization and association with the world. This issue of contextualization, you know, becomes less important and this hands-on experience on the language gets a crucial point of communication. The task development needs to be different. For instance, it must have some emotional and cognitive orientations. For instance, you want your language users to test their language in a particular situation.

You can ask them to go into an open place and use the language just to see if it deteriorates or improves testing conditions. We know that something that isn't academically assisted can affect the response value of the test-taker negatively. What is happening in the Internet-based testing? The space is too liberal for the test-taker. For example, the time-out and break are normal. Yet, the tester and the test's administration want the test-taker to establish his/her response appropriately. A social condition for such appropriateness is significantly true for a change agent in the assessment guidelines for a classroom and its groups of learners.

We know the language in office is different from a classroom because of its consistent linguistic behaviour. When someone wants to interact in a field styled environment, it has its own functions to perform. This task with its placement at a time and place with some people has gotten good measure of a decision in the assessment.

We can use all of these grammatical test items for the development of the test's tasks containing the test taker's emotions and cognitions. I think that's missing in some international tests. If we have gotten classroom learning opportunities, we can exercise the same in the conditions. We can also push this learning of the testing off the corridors of physical realizations of formalization. That's how, you know, the language that is simulated in classroom gets its real perception and practice at

—

another site, the same room of interaction. If we can ask our class fellows or test takers to finish a task by visiting a place, spending some moments, talking with some gentlemen and ladies and reporting the same back to the tester as an examiner in the language in question—and we have those grammatical objectives in mind that the person who has been asked to visit the place for such talks is communicating (grammatically) and using right inflections, right derivatives, right order, right sense of completions, right words and combinations and right phonemic arrangements—this is how the syntactic justification can have good semantic match with the mind of the user as a test taker and the knowledge of the expert.

If I can sum up, I've just introduced you to some ways of grammatical tests today. Most grammatical tests are decontextualized. Therefore, they lack the test taker's abilities' measure. They are used to predict their accuracies towards the language use. The test begins with MCQs. It is quite decontextualized and, I think, ends with some combinations and sense completion. The beginning is more decontextualized and the ending, I think, is more contextual. And to get out of this issue of the continuum of the decontextualization or contextualization, Alderson is of a view that the models of languages in Europe are a good help for the description and development of test item as a task practice. This task practice needs emotional and cognitive abilities of the test takers.

They can be used for language predictability. If someone wants to be assessed for their grammatical accuracy, he or she can be asked to visit a place, have a word with some gentlemen and ladies and then report the same back to this examiner at the place of learning. And this examiner as a tester can have some accuracy control. For instance, this reporter who has gone to the place and come up with some information is likely to report back in some fixed inflections and derivations. He or she can also develop the same into some complete senses with no issues in order. This syntactic regularization is going to match with their semantic development. And the same is adopted in linguistic knowledge of the expert. Thank you.

-

This error is not the fossilization of the language user; it is rather a limitation of the language corpus and its permutations extended to the human user for a creative stretch of the expression of or about or through language or a language. A red highlighted part that we see in a document match in MS-Word establishes as an error.

-

CLASSROOM TALK 14

Vocabulary testing is a fascinating job of a tester of English language education at the moment. He/she really wants the test takers to enjoy the active and passive vocabulary items and their expressions and the language they use. There are a number of ways of getting to the same process of language testing. Before we go to this, we have to make lots of decisions to place these items as test items. For instance, vocabulary in itself has its own operational definition in language use, we know. We also want to use this test as a tool for communication. It is the application of the communicative practices of all practitioners of the language including the testers, the test takers and the administrators.

If we go back to its operational definition, we know that it contains lots of words. For instance, the same words that take certain shape into the minds of the users have their physical reality. One of the realities of such words' mental processing is phonological, aerodynamic and physiological. That is this aerodynamic part of the user is perceived back in the brain and the body by joining in the air and processing it throughout the body before it is pushed to the environment of the same in such a way that makes meaning. I think that gives this vocabulary, right? We also refer to this as a vocabulary. I think it doesn't turn into this vocabulary or word bank, but it does contribute to the definition. The physiological one is developing in the body of the kid and this aerodynamic one is perceived in sense and pushed into the body for its development. The same has gotten some sounds that make meanings to the mind of the user. I hope that makes sense here. I want you to pay attention to this phonological one before we go to the other part of the same vocabulary building operation in the mind of the user. You have gotten lots of opportunities. You can use those phonemes, syllables and word stresses to help the user recognize their sense

—

developments for the continuation of a specific language in a specific domain of communication. For instance, the word *read* is a bit easy. And it has gotten some phonological interpretations with the body that receives it. The body that wants to produce or reproduce the same has to go through some physiological and neurological developments before the same production is placed at a particular point of the structure of the language.

This phonological way can also be tested for the promotion of the vocabulary. Its second physical reality out of those many realities of this mental processing of its operation is to do with morphological development. Let's say phonological one passes through some morph and morphemic development. That means that these phonological, phonemic, phonic and sound patterns begin clustering themselves to certain points of repetitions in one's language. Now, when all the *-tion* exists together, in most cases in the language of English, we know that that is quite English and that comes at the end of a complete sentence. The same cluster repeats frequently in this structure. There are suffixes and prefixes in English, but when we go to our own language, we know that there are many infixes. For instance, Urdu has them. The complexity of linguistic information (phonological and morphological) is higher than the two at the ends of an English word.

Please remember there are some infixing at the level of phonological generation; but I think that isn't the right time of the discussion in question. If I go back to the same point of the secondary language, the phonological development tends to morpho development. That is, the same cluster of certain sounds in phonemes repeats, so sense generation and its match with the world and the people in contact. Affixation helps establish word classes. This combination of certain morphs can also be translated or transformed into certain word class. For instance, this is a verb or that a noun. They are used at different places of the structure. We can't test English structures here. We have tried our best to aim a test to test our vocabulary items, right; but we know that the same test has this complex arrangement of the language, which is inseparable from the mind of the user. This separability is administrative, not linguistic—not linguistically trained component of English testing. There are many morphemes, right? We know that there are some bound morphemes and some free ones. Those free morphemes can have certain word classes if they are used with some bound ones. The bound morphemes can't exist independently in the language of English. This first reality, the phonological, has turned into the morphological.

And the morphological gives us the second reality, that is, word classification. We also know this phonological, with this phono-morphological development of the language of English, has its structural value. If this sound with this neighbouring one occurs and

—

recurs at a point, the phonological mind is capable of generating some sense of mutually shared meaning.

Mutually shared means that something happens intrapersonally can also be exteriorized, right? If something is intrapersonal, it becomes a part of interpersonal communication that needs a syntactic development. A phonological to morphological arrangement gains some importance of meaning at the level of its generation. The word, say, *fast*, has its phonological awareness. At the same time, it has its more phonological reproduction that has its word classification. Its appreciation is usually seen if its reproduction is used correctly—appropriately rather. And if you say appropriately rather, it means that the person is going to a syntactic development of the user's mind, which is more physical reality. That's your third reality of those many realities of vocabulary item's definitional operation.

The first reality is phonological. The second is morphological. The third is syntactic. And they combine together to give us a language. This is consistently called as lexical development. This lexical development of vocabulary has this complication that needs this separation and specification. Please remember the same lexical development has a semantic value. That is, syntactic, which is physical, has this mental processing, which is rather soft, virtual, personal and public.

For instance, if the word that is tested has its referential value, it is easy. We know that we have that referential value. It's usually defined with a denotative one. That is, the word meaning is available in your good dictionary. If this is tested here, it has its meaning in the dictionary. That's easy. Or that the meaning is in your syllabus. It means the syllabus that contains lots of reading texts and talks and listening comprehensions and many other ways of reception and meaning making ways. Please go back to my first talks on the site. That's denotative. There are many connotative meanings. For instance, there are some social meanings. There are some affective meanings. I think I should use the word *red* here because I have just gone through a website with all color of red. With lots of people from a specific race here. This needs its representation in today's world. I would like to share the web too. The web has gotten the title RED. And it has gotten lots of products to promote across the world. The same RED which has its referential meaning in the dictionary is a matter of colour. And its color has its chemical generation by now. And the same can be taken and tested by now. And there are many ways of doing. For instance, you can show the same colours to the user and ask him or her to match it with *red* and match it with the phonemic organization /red/. Or that can also be made more complex by placing scrambling phonological words or phonological components or parts. And you can also show the same word. And you can ask him or her to check if the word has this meaning. And if the one who can read the word...I think, he or she can also go to the

—

fourth physical reality of those many realities of this processing of the operation of the term vocabulary item, that is graphemic. That means that the word written has this graphemic orientation.

That red has R-E-D. And they have gotten in their certain shape. They do change even if they have their flexibility of stylistic variations at personal and public levels. There are four physical realities: phonological, morphological, syntactic and graphemic. Please remember the fifth one. I think I should not consider this physical one; but that is something in the user's mind. It's materialized into the physical reality, that is, semantic reality or semantic development. It's referential one. It has denotative and connotative. The denotative has the meaning in the dictionary. Social means, say, *red* refers to love. For instance. Affective meaning of red refers to anger. Or *red* can also refer to that social meaning. That meaning can also refer back to this referential one. And then another meaning that can also be tested... That is metaphorical, which is an extension of the language. It can also help the user enjoy some literary section of life. The semantic development of this lexical orientation of the vocabulary item has gotten three to four components: referential, social, affective and metaphorical. You have gotten lots of opportunities of placing your objectives before you develop the same into certain test items. The same semantic development should take you to those dichotomies of meaning. For instance, hyponymic development. We know that the concept generated can have some subtopics with some concepts in contact. For instance, *red* which is a colour in an image back in the mind, it's also in a body. If something bad happens to the skin, say, and if I cut the skin for some reason, it needs to be red. We can also show the same colour using the same human body. His conceptual development begins with the worded *colour*. It is both hyponymic and hypernymic orientation. It will show help test the test the word. Before you go to test the word, you have to go through the process of this word formation which is physical and mental. The other dichotomy is synonymous development that the words can have some synonymous meaning. We know that English doesn't have those exact synonyms, but there are many synonymous ways. We can replace some words. We can use certain phrases. There are many synonyms. MS Word offers this feature. We can use it.

If we think that the word repeats and that the word doesn't sound good in a particular environment of the words of that sentence, it can be replaced with MS Office. This way of doing is more synonymous if not exact synonymy. For instance, red can be replaced with scarlet. We can say that red has a synonymous scarlet; but we know that they are different. We can use such words. This hyponymic development can be turned into a synonymic one. There are many polysemic derivatives or developments. For instance, *plain* has gotten different lexical entries in dictionaries. I don't know if *red* has had different dictionary entries. A polysemic expression has many meanings and

—

entries in a dictionary. It does depend upon the lexicographer. Sometimes he/she has one entry for a meaning. Sometimes he/she wants to add multiple words of the same graphemic orientation. We can also say if *plain* has this meaning of this *plain* in the world and the sentence or in this occurrence of this listening comprehension has the meaning of the same word *plain* in that specific utterance of that specific speaker. We can say that they are different even if they enjoy the same phonological generation and similarity; but the semantic sameness should be questioned and discussed separately. It can also help test such a word.

There are many homonymic words in the language. The word *read* can also be homonymized. Graphically, they are different; but phonologically it is homonymic. We refer such combinations to as homophones if not homographs. Sometimes, we test just homophones. Sometimes, we want to test homographs. Sometimes, we want to do both carefully.

Let me summarize. I started off with vocabulary item's importance in human life. It begins with a development as a reality. Something that develops in the mind is more virtual—not physical—has its aerodynamic and physiological development. The aerodynamic part of it is in the environment. It's also part of a body that helps breathe in and out and that the same is a process for mental processing brain development and heart circulation across and all in all the body. The same phonological reality turns into morphological one then into some syntactic one and then into a graphemic one. The syntactic development is central to the development of the brain and its activities for the mental processing of words and meaning conclusive contests and changing requirements for interactions and negotiations.

All of them can help establish this mental development which is referred to as semantic development. We know this mental development has a number of ways to defining that the referential which is denotated and connotative and social that is more community developing and then affective. That is more emotion wise and metaphorical which is rather literary. The semantic development leads us to semantic dichotomies from hyponymic to synonymic to polysemic to antonymic. Sometimes these are binary opposites. Sometimes they are non-binary. The morphed white and red are improved in intensity, with their suffixes, in English. There is also a fine point of consistent gradability in the language of English. It's also in the language of Urdu, Pashto, Sindhi, Balochi and many other languages of the country that we live in.

How can we test all of these sections of life for vocabulary?

I think there are many ways. We can develop some MCQs. We can have a cloze test with some blanks in a passage in a systematic order. We can give them a passage of audio/video and reading. We can also ask them to follow their syllabus containing

—

some texts and the same words and phrases can be used differently on the day of the test. Swedish has a phonological similarity with the language of English because of its speakers, I guess. They are considered more proficient than the ones of those who live here with us in Pakistan. Marklund *et al* (2018) want us to test the development of a vocabulary with early expressive phonological skills. That is, physiological developments in and across human bodies should have some early interventions. For instance, the babies don't want to pay attention to their non-obstruent sections of life. They are quite good at doing so. They want to make some obstructions. That is, they want to use their articulators for the production of their awareness to the environment. They want to use their lips, tongue and teeth. If they have begun sensing and want to use their gums, they want to do lots of things in early years. Such interventions—sometimes such parental interventions—can help them cluster their consonants together, with some syllabic structural developments. These syllabic structural developments will have some vowels, non-obstruent way, which is referred to as sonority. For such kids, that isn't the sonority; for us, that is. For them, the sonority is meaningless. Consonants, obstruction, their physiological labor that is driven through their physiological development and aerodynamic utilization are a significantly true in the literature of early language acquisition.

We should not (explicitly) teach English as a second language. We should teach our kids the language the way it's acquired today. We can intervene. We can have a parental intervention for early expressive skills that this part of the sense needs a stress and this one needs a syllabification and that can also gain emotional and cognitive development. The consonants, vowels and their combinations can also have such vocabulary development, testing, procedures' lead in. If we go to adult education, we know that vocabulary has gotten its active and passive parts. Words are frequently and infrequently used. Sometimes you should test their active vocabulary items. Sometimes you should test their passive vocabulary item. They are received but not produced. Active vocabulary items are those that are received and produced together. And there must be a way of taking their memories to the point of retrieval, a schematic part of it.

Let me summarize. This vocabulary has its early education acquisition and adult education. Early education has its consonant development in the presence of vowel sounds. It's enriched with positive parental interventions. Adult language proficiency needs its division into active and passive vocabulary items. The brain can help reproduce all words if they are provided with some impetus. Some vocabulary items are more specialization-based, some are daily life-based, some are foreign and some local. You have to decide what kind of vocabulary items you want to share with your test takers. If you want them to perform on some specialization grounds or if you want

—

to do for daily life practices or if you want them to enjoy foreign environment of languages.

For instance, in the country of Pakistan, the foreign environment of the languages is the UK or the US or Europe or America or Australia or Canada in a normal linguistic condition that lacks interpersonal communication. I don't know if this foreign environment is the foreignness. We consider it here but there are many negative points that, I think, should not be discussed. There are some local ones. Local refers to the community in contact can be tested. This immediate long orientation can be tested this way.

Heaton writes in the book:

With a careful selection or sampling of lexical items for inclusion in a test for a most exciting and exacting task. Many of the more traditional types of vocabulary tests are designed in such a way that they test their knowledge of words which frequently found in many English textbooks are rarely used in ordinary speech. The first task for the writer of a vocabulary test is to determine the degree to which he or she wishes to concentrate on testing the student's active or passive vocabulary. The next task is to decide whether the lexical items in the test should be taken from the spoken or the written language. Selection of the vocabulary can thus be thought of as falling into the following rough divisions according to the four language skills: listening, reading, speaking and writing.

-

The first reality is phonological. The second is morphological. The third is syntactic. They combine together to give us a language. This is consistently called lexical development. This lexical development of vocabulary has this complication that needs this separation and specification. Please remember the same lexical development has a semantic value. That is, syntactic, which is physical, has this mental processing, which is rather soft, virtual, personal and public.

-

CLASSROOM TALK 15

The last handout's procedure containing lots of points along with the boy's picture in a cap in an angry mood should help you go back to it and observe again.

Bachman (1991) suggests the operation in three ways. Firstly, the language needs a theoretical complexities' enrichment in language test operation. This operation of the test has many components. If one wants to test English language, he or she is likely to come up with a unified thought of language proficiency of the user even if he or she knows that this particular score that the user as a test taker has gained is about that particular situation of that particular content or the item response to the test item at a particular time and particular moment. There are many variables that the tester is aware of; but he or she, I think, is likely to come and stay consistent with a unified score for global needs. That is quite challenging. At the same time, I think the theoretical complexity contains test taker's characteristics. The test developed must have the test taker's characteristics. I think the same part of Bachman's ways should lead you to the appreciation of language test authenticity containing situational and interactional ones. We know the situational one that language tester places some specific position and its intractability is to do with a test item involvement into the abilities of the test taker. We can assume that this interactional authenticity has this motivational part of the test taker's abilities' involvement. If the test taker has gotten a good ability and a particular skill can be used for this involvement of the test item's completion for a score, the need of its representation or generalizability should be highlighted. The representation means that need is concurrent. The validity and generalizability suggest a predictive validity. The concurrent means that the same time used has gotten this score predictive. The validity means that the score obtained that moment can be used for future practices of the test taker in this certain particular context of language use. Secondly, the second way of Bachman is a suggestion of a

—

statistical measure of the performance of the test taker. It is tricky. On the one way, it is found dealing with theoretical complexity containing test-taker's characteristics for the abilities' involvement into the achievements and completion of the test items. The other way is about its calculation. I think, that is quite rater based, evaluator based and/or reader based. The reader is the one who reads the response and the rater is the one who rates it. The reader can also become a (perfect) rater. Both of them should be perceived as good evaluators and examiners. In Pakistani context, we consider them as examiners. What is the statistical measure? We know that the statistical measure should lead us to good readings of the reliability. This reliability is about the test score's evenness. This test score evenness should be understood in a sense that if the test taker scores 7 out of the 9 benchmarks in the IELTS should get the same if he or she doesn't improve or learn the language. He or she is using the language; but if the same test is reconducted, the test taker should get the same 7 out of 9. If he or she gets the same score, the reliability is acceptably high. This indication of the performance of the rater is called as rater's reliability. I know that the score is shared between the rater and the test taker. I think we have moved a bit ahead because the tester needs to be a rater or a reader. The 7 should be shared between the test taker and the rater. The rater who suggests to define this numerical value out of the benchmark 9, excellently defined and tested in the IELTS testing system, for the same candidate at two different times with no obvious learning practices is consistently true in their performance. This performance is considered as intra-rater performance. It is quite happily consistent. The inter-rater consistency is derived valid and interpreted high when the same 7 is given by two or more different human raters for the same candidate. A change in the score at very moment and its interpretability happens to cause a depreciative value of both human consistencies. Therefore, good education and training are a central element of a testing through/with a language. The language's inevitability must not fall into a human vulnerability. Many computing devices and software packages are available in the market. They should be used for a global need and a satisfactory standard.

Finally, the third way is communicative approach. I know the communicative approach has been successful in the recent past. Let us see if the same can be followed here. We know that the language in question is a library language in most Asian countries that we can access to. The library language is not that communicative controversially. This controversy means that the one who uses this language for silent practices has to come up with controversial phonic sounds. That is, the reading has to initiate listening and speaking in a foreign culture. Or the language must not be spoken but read and written in school and college. This is what is true if you have noticed in the University of Karachi. The students hesitate to participate in the teaching time of the language in the classroom. I tried to do it in the first five years in those small and some safe rooms of the department of English. They hardly did it.

—

There are plenty of reasons, I guess. The students do not want to equate themselves; rather they need the information by a teacher. Or they cannot take turns neatly in their early years of university learning. What is important is that they should be given this opportunity in the K-12 so that they can voice and improve the discourse here. We have to get out of this loving Iranian, respecting Arabic, mysterious guru Indian and religious practices and cultures of reception and should look at this approach placed in this global wisdom of human communication.

I am going to give you four tasks from John Wright's (2003). I would like you to use them as teaching plans. You may use the same teaching plans at home and in friendship and see if you can teach and test and if both of them can complement each other. It will have a good positive backwash. We will have to begin with Bachman's introduced theoretical complexities including the test taker's characteristics and continue with the statistical operation:

- that we should have some rubrics,
- that we should have some descriptors,
- that we should have some statistical measures,
- that we should have some intra-rater and inter-rater consistencies in the context of the reliabilities of test score evenness and
- that we should check the communicative approach beyond library language practices.

The second part of language test operation contains the most crucial section of language testing; it is language test validity positively contributing to the communicative approach. The validity is its acceptance and promotion. For instance, if the tester wants to test the skill of listening, he or she should design a test item as an input aimed to measure the same statistically, communicatively and theoretically; and if someone wants to test writing of an undergrad, enrolled with the programme of astrology or astrophysics, he or she should be given some relevant content developmental impetus. That is, something about the same field or something about the same test taker's intractability content is required or presented for the language's production and growth. It is likely to increase the authenticity's interaction caused by the involvement of the test taker's abilities into the test item and its completion. The completion means that this input will get an output as a response. If the test taker is English literature's student asked to write about physical sciences, he/she is unlikely to respond appropriately. This input will fail even if it is one of the best ones in the world because of this lack of the validity and/or interactional involvement of the abilities of the test takers into the test item.

-

I think that the same issue should be labelled with face validity that something needs to be interactional and that something that the test taker has to accept frankly before he or she takes it up successfully. Please remember the third way should lead us to the content validity. The content means the language use must have its positive relationship with language tests. Similarly, the teaching should lead to the testing; or the testing should go back to good teaching practices; or syllabus objectives must be achieved through such tests items' objectives and displays to the test taker's responses' materialization and successful measurement for a decision. Another way of language validity is referred to as construct validity. The theoretical display is an important part of content generation and its interaction with and across the respondent and the writer.

What kind of tests do we need to design?

- We have to look at those languages' skills.
- We have to look at those knowledge skills.
- We have to couple those knowledge skills with language ones.

We know that language skills are hardly quantifiable if the knowledge is missing. We must set some knowledge before we go to do it for language skills. We know that the one who wants to learn the language imitates. I was discussing the same in the recent past that the one who did it would imitate and then learn processes and would create and industrialize. This imitation develops into some learning practices and their complexities because they are dependent upon lots of variables including theoretical and statistical. The language has its creative part that no language user can reproduce the same response in a subjective order even if it's objectively verified. It has its different psychological operation back in the test-taker's mind. That's why it's referred to as creative language. Strategic competence is central to (creative) writing. The creativity can be measured when the test takers is given an opportunity of writing many words with their own decisions and strategic competencies.

The same language creativity can be industrialized. The industrialized means that the language produced is always public. If someone has written a piece of paper, it can be shown to lots of people, addressed and unaddressed audience.

If I can summarize today's talk. I started off with the language test operation. I tried to use Bachman's three ways: theoretical complexities, statistical measure and communicative approach. I really love to go with two different scales: language validity and language reliability. The validity has face, content, construct concurrent and predictive any many other forms. It is usually achieved when the test item's language gets high authenticity containing situational and interactional. The

—

situational is more test taker's and test administrator's dependent and the interactional one is more test taker's abilities' dependent. The reliability has a rule to perform for language test score's evenness, with (perfect) raters for the inter- and intra-rating consistencies.

-

I know the communicative approach has been successful in the recent past. Let us see if the same can be followed here. We know that the language in question is a library language in most Asian countries that we can access to. The library language is not that communicative controversially. This controversy means that the one who uses this language for silent practices has to come up with controversial phonic sounds. That is, the reading has to initiate listening and speaking in a foreign culture. Or the language must not be spoken but read and written in school and college.

-

CLASSROOM TALK 16

It's the last session of the course. I would like to talk about the English language programme's evaluation's section's description. The evaluation is too huge to be discussed in the course at the moment, but we can touch upon the relevant part of it containing some language testing practices. We know that the test in question has its standardized ways and have tried our best to have some agreements upon some procedures. The testing is improved. This agreement can lead us to some standard value. Whatever happens in the country of Pakistan is as same as in any other country. We know that the places and the people are different and have to adapt ourselves. A good conceptual congruence is central to the success.

There are three simple ways to the improvement of the test procedures' evaluation. We must determine the significance of a test item before it is planned to evaluate the test taker.

We must measure the test item difficulty for its separation from other sets of the items. There is an opportunity of pilot studies that I remember. A pilot study or a small-scale study can be completed before such a set of test items is sent to the right test takers. Before we plan to test or assess the test takers, we should see if the test items are also right. Since you have to measure the test item difficulty, the test item can be too easy or too difficult if it isn't made good for the testing procedures for a good statistical control over the knowledge and its users. However, I have this conviction that a good user of the language must get a good score irrespective of this control for the best fits and the outliers. I think they should be considered as the outliers in both of these extremes and such test items may be considered into this concept developing notion of the performance and the performer. We have to get out of them and go to the middle part of the test aimed to test a group, equally difficult

—
for all the test takers. This needs the right test-takers for a test design.

Theoretically, we know that when we get some responses, the same theoretical similarity of convergence of the item difficulty will be displayed on a continuum of some good spread. That is, that these outliers—if removed from that too easy or too difficult type of test item from the set of the test items—this common similarity that you see right at the centre of this difficulty range or challenging range, will have some different responses. Those responses can be spread on a scale between/of 10 to 100; but it's really difficult to have a test taker scoring 100 out of 100 or 9 out of 9. It's too difficult. If it happens, we have to question if the test was really too easy to the test taker.

What is suggested here that the test has its denotative meaning if it is used to assess public achievements. The University of Karachi's entry/placement tests have not been used appropriately, yet the students want to perform in the oddities of the facilities and faculties. I remember an outcry of the public and the university's resentment over the monetary crises. More interestingly, the process for the MPhil and PhD candidacies in English here is a political monopoly of the governing system establishing its thematic wisdom of worldliness and rewards and punishments in English/European influenced lawcourts for a state operation, not the community of scholars. The lawcourts have human themes drawn from the pre-renaissance, Victorian lifestyles for the nation too underdeveloped. This view is not intended to resist the state but its language means for a bad practice in the country if it not drifted away from the course of the Language Assessment and Programme Evaluation. This opportunity that I have created for the close of the semester teaching should be utilized for the separation of English literature and English linguistics. The former is more universally accepted notion for a propagation of human themes and world order; however, the latter is aimed to improve life sciences. One is fundamentally history governed texts; the other is a continuous developing language and its technology. One wants to stay; the other wants to change. Linguistics isn't a theoretical resistance but a part of technological swiftness today; the literary values are human controlling if not dictating. World Englishes and Multilingual Education are good examples of the same institution. The democratic value observed for a test design is hardly seen in the testing procedures of the candidacies. The prospective candidates are asked to write research proposals of 2000-5000 words on a day of the testing. It is a mere reproduction. Nowhere is it seen for a research work. The whole construct of the institutionalization of language testing knowledge collapses. The result is the candidates selected cannot finish the studies in the department. Even if they do so, by any means, the suspicious methodology needs its validity. The university's professors who have finished their studies in English are unable to use the language, unfortunately. The recent artificial intelligence is a potential reason for a justification

—

for crime prevention in university education and its better utilization for language support. This bad test design may be interpreted in the contexts of political improvisation of some human groups for the representation of bad governance of public institutions including power supply units, local government security personnel, health, road communication, human communication, national data support centres and aviation installations, and primary and secondary educational schools for public propagation of deviant statehood. As a result, the state does not enjoy a good system of craftsmanship if compared with other developed countries, but it falls back into unwanted tribal practices and incorrect knowledge development for skills and manpower for a conceived modern peoplehood of Pakistan. On the other hand, the global woman's positive discrimination seems to have influenced the test designs negatively here too. The skills and support the country needs are made possible if the state rejects rampant corruption and intends to follow the standards. Every good graduate struggles here. It is not the academic scroll; it is the learning that helps the nation and the national. They must coexist.

More to this degradation of the influence of the university education, the students do not have their face identity. They consider it a matter of Islamic values. However, they have it in their national cards and move freely. Their thumbs impressions and hand signatures are not verifiable. The administration is too weak to determine the testing's administration for a such mass value of the testing's impact upon the world and its nation-state operation. Power-cuts and bad rooms, bad acoustics and bad instructions add to it. The test-taker's characteristics are the third way to the improvement of the procedures. Let's come back to the normal testing activities.

The test taker must be given some challenges. They should be taken up successfully. Those responses of the test taker should be interpreted.

If I can summarize what I'm trying to say here that the test item's difficulty or easiness is measured before it's sent to the test taker. With some pilot studies/small scale studies, its measurement can be carried out. It needs some significant funds and time and people. It doesn't happen in Pakistan, I know; but, as a professional, we should see its important value addition to the test's validity. The second part of it is the measurement of the rater's severity. It means that the reader has gotten their own preconceived notions and experiences even if they are too trained. A classroom teacher should be trained to test too. I think that the best teacher can be trained in a 30-hour workshop. The rater's severity is problematic. A rating scale or a rubric may be utilized if the conditions are favourable. The perceptions for different numerical values are problematic too. This severity should be reported separately for a more sameness of a score. A good testing team delays the test taker's interpretation to get to a good decision for the programme and its impact upon the community, the nation

—
and the country. I think this delay is really significant. I know that administration is enhanced every moment. It is perceived efficient if it does something quickly. We know that it can do quickly if it has gotten a good procedure. This delay needs some steps. The two are the test item's difficulty and the rate of severity measure.

This is how numerically good or bad test taker is to the test item. The test taker is good or bad to the test item, not to the test or tester. Sometimes, the tester can't read efficiently. Or they are provided with hundreds of responses in a week. They have to hurry up to meet the deadlines and make mistakes, not because they want to, but because of the human constraints. If you can look at the same sheet, you will see that this improvement that I want you to see begins from light blue and ends at dark green. That light blue is an indication of the capability of the testing team for the numerical values collection. All of these three measurements should have some numerical evaluation. If they are provided with such values, that is, how good or how bad or how severe, how simple or how easy, how difficult, from, say, 0 to 10 or 1 to 10. To collect those numerical values, let's put them into some data. We know that putting such a person to the values has its procedure too that I can't discuss. And the same should be transformed into some data. Again, that needs its procedure. It's more statistical and I want you to do some statistical course. This is how you can move from light blue to dark blue. And this dark blue needs some positive, positive, positive numerical significance. That is, we should have good frequency development. We should have good range and good central tendency. The points 1 and 3 on the sheet go together here.

What is the central tendency?

What is the range?

And what is the frequency development?

The frequency is a statistical fact that if you have gotten 10 responses and 7 of them have been labelled with 6 out of 10, the most frequent score 6 is used to determine the replies to the questions.

I don't know if I can continue discussing it. Let me rephrase.

If a rater is provided with 10 extended written responses, he or she tries to check grammatically accuracy, content development, train of thought, writing conventions if necessary and writer's clarity of involvement. A conceptual development of the rater's inclination towards the rating must be observed and reported back to the team for further improvement and public research. A good transparency for the testing is equally important for public wealth. Please go back to Heaton's Chapter 11. Have the testing's research funded. If you can have your research plan funded, you have gone

—

all yellow here. The funds are always fascinating. You can buy books, get some tools and employ some team members. The positive backwash effect is celebrated. The points 1 and 3 will have a stronger relationship. You can get out of your place and discuss with some competitive team members and heads of different organizations. You write it and want to publish. If you can do all of these, you can have a touch of this light green. I think this green part is not bad even if you can't move to the dark green because the dark one needs some authority. If someone wants to improve, the publication is an excellent way. The change is impossible unless if the government, the learning community, the organization, the people and the culture are brought into a favourable condition. The government can read and make decisions, can't it? We stop here. This is the last session.

-

Linguistics isn't a theoretical resistance but a part of technological swiftness today; the literary values are controlling, human controlling if not dictating. World Englishes and Multilingual Education are good examples of the same institution. The democratic value observed for a test design is hardly seen in the testing procedures of the candidacies. The prospective candidates are asked to write research proposals of 2000-5000 words on a day of the testing. It is a mere reproduction.

ASSESSING ENGLISH LANGUAGE TESTING

NATURAL STREAM OF THOUGHT INTO SPEECH AND SPOKEN PERSONALIZATION

- Reading aloud
- Reduced redundancy but varying level of utterance frozen forms of sentential repetition
- Spoken words strung to an utterance and its combination for few to more words per minute (say, 50-300 words) for a situational linguistic readiness

HISTORICAL SUPPORT UNITS

- Theoretical supported operationalization
- Cognition and metacognition
- Competencies/tasks
- British test developers using academically interfered/improved English language acquisition/learning
- American test developers using psychologically interfered/ improved English language acquisition/learning

STRUCTURING THE NORM OF LANGUAGE TESTING

- The item as a task (difficulty doing projection of task links)
- Time allocation (linguistic reception/response)
- Place allocation (test-developers/test-takers/language users)
- Content allocation (proficiency/diagnostic/achieving/achieving/placing)

-

US DEPARTMENT OF STATE- FSI

The mission of the Foreign Service Institute (FSI) is to deliver world-class diplomatic training and provide the career-long learning opportunities that U.S. government foreign affairs professionals need in order to excel in today's global arena, advance U.S. foreign policy and deliver on behalf of the American people. To fulfill these critical responsibilities, the personnel we serve require a unique combination of specialized expertise, operational skill sets, and a broad capacity for innovative and visionary leadership. FSI training and resources promote substantive, regional, and linguistic expertise, leadership finesse, personal resilience and problem-solving.

OLDSPEAKINGFOLKS

Situation + Participants + Ends + Act sequence + Key +I instrumentalities + Norms
+Genre

TODAY'S ORDERLINESS

- Genetic relationship to speech physiological and psychological fits for (spoken) speech production
- Political governance of grammaticality of standard varieties of English
- Native-native, native-nonnative and nonnative-nonnative interlocutors' potential as language testers and test-test takers
- Accentuation
- Speech rate
- Topic and tail

LANGUAGE TESTING COMMUNICATION AND COLLABORATION

- Topic and tail
- Theoretical supported operationalization
- Language appropriateness
- Impetus/intrinsic/extrinsic motivation
- Communicative relationships
- Friendliness
- Interlocutor or/and test-developer
- Generalizability

—

- Static/dynamic picture incentivized
- Audio directed
- Personal relationship focusing
- Impersonal human assisted
- Open and structured
- Pronouncing
- Repeating
- Reacting phrasing
- Personalizing
- Accentuating
- Interviewing
- Pairing

DESIGN THE ITEM AS A TASK

- Narration for the past temporal relationship and grammatical inflection support for the language
- Description for the information structure of the mind and the mouth
- Exemplification for the interaction
- Comparison for thinking patterning enhancement
- Evaluation for personalization and professionalization
- Regularly sight assisted mental support and (developing) schemata of colour, line, and shape with orthographic knowledge drawn to be written and/or typed semiotic occurrence relationships of syncretic, synchronous and similar meaning sets for the learning and learned.
- Vocables support
- Speaking protocols
- Human language knowledge reception and retention
- Lexemes as vocabulary items

EYE THE WORD

Varying level of comprehension resulted by slow and lazy eye movement because of demoralizing, demotivating, too much cognitive burdening and too higher for meta-cognition or exciting and fast eye movement because of information seeking likeness, problem-solving encouragement, its translation into speech impetus and brain and body well-being.

—

This orthographic knowledge's reception is an indication of local and global comprehension and reaction to it. The psychometric measure may be used to identify the reader's ability for a good degree of comprehension of a written (non)academic and/or (non) literary text's length in a timed place of assessment for diagnosis, proficiency, achievement, aptitude and placement.

TESTING TREND

- Reading aloud
- Silent reading
- Differentiating
- Relating
- Topicalizing
- Text specifying
- Meaning sensing (enunciating)
- Meaning forming (pronouncing)
- Meaning integrating (speaking the same)
- Meaning finding (fill-in)
- Meaning generating (extended written responses)
- Meaning negotiating (writer-reader and reader-reader overlap producing or rejecting)
- Comprehension accuracy range
- Testing realizations into
- Fill-in
- Ordering
- Matching
- Yes/no
- Choosing
- Cloze
- Partial writing

READING CHARACTERISTICS

- Intent involvement
- (Sight) assisted effort
- Systemicity
- Self-direction
- Problem-solving approach
- Rules and skills (conventionality)

-

- Friendliness
- Meaning
- Mental alertness
- Strategic support

VISUALS, SYMBOLS AND WRITTEN EXPRESSION AND TASK

- Cognitive (Psychic arrangement of information for a TASK BAR)
- Cultural (Rather socially placed, performed and posited VISUAL REGULARIZING TASK BAR)
- Situational (Communicative spirited conscious streamline for cultural cognition of symbolic exteriorization into human rhythm of WORDED AND PHRASED TASK BAR.
- Educational A rigour of practice of the controlled classroom discourse of STANDARDIZING TASK BAR.
- Professional INSTITUTIONAL WRITING TASK BAR
- Formal A slow mechanism of draft, edits, proofread and public acceptance of writing as a PROCESS TASK BAR.
- Liberal A possible pursuit of a stylistic, genre specific and/or modified expression of artistic value.

HISTORICAL WRITING TESTING DESIGNS

- Theoretically driven curriculum support
- Positive local and/or global backwash shifted
- Direct writing
- Indirect writing matched
- Language production on paper or screen
- For English words
- Sentences
- Paragraphs
- Topical extended written responses
- Tells or transforms of oneself, one group, one genre, one identical linguistic relationship and one or many facts and foci in a timed or untimed year of learning, performing and working.

-

TEMPORAL TYPES OF WRITING TEST ARRANGEMENT

- Quick response
- Timed response
- Delayed response (portfolio writing)
- Untimed response (reflective piece)

WRITING TEST-TAKER'S CHARACTERISTICS

- Grammatical accuracy
- Thinking control and development
- Topical knowledge
- Language knowledge

WRITING TEST-TAKER'S LIMITATIONS

- Legibility
- Readability

A WRITER TENDS TO BE

- Personalizing or/and distance creating language choice
- Language following
- Connecting
- Developing
- Creating
- Publishing

LANGUAGE TEST, TEST AND TEST-TAKER

- Reliability
- Test items' words and topics
- Rater's consistent display of impressionistic or analytic scoring, valuing and interpreting
- Test administration
- Validity
- Meaning richness, global connection and authentically correct

-

- Addressing
- Engaging
- Situation occurring
- Problem solving
- Reporting
- Life practicing

LANGUAGE TESTER ACTS

- Impressionistic
- Analytic

LET'S THINK OF

- AI subjective nature of language (nonhuman support)
- Physical support (human fits)
- Grapheme controversy for English (human semiotic constraints)
- Morphosyntactic support (human semiotic improvement for sense structure)
- Linear publishing and printing (human non-customized addressed text)
- Nonlinear publishing and printing screening (human customization)

TEST DEVELOPMENT INITIATION

- Test taker's information (sex, ethnicity, language, education, age and goals)
- Language difficulty
- Learning
- Learning and performing
- Learning and professionalizing
- Naturalizing

-

A paper to read.

-

A RATER'S PERSPECTIVE

Why does an educational team need a rater? Two most likely reasons for this decision are the understanding of the language's complexity and its use among the speech community placed in for a human expression and the utilization of the knowledge for its measurement. Once the syntactic complexity is established, the practices of the user should acquire other educational habits for their growth and liberty. What are they? How can they be determined? Why is a rating scale developed for an academic reason? This study offers an action study about some written responses of a course of World Englishes taught in a programme of English at the department of English of the University of Karachi.

SYNTACTIC COMPLEXITY

Syntactic complexity of a human language is a natural phenomenon to witness in the human user's production. Although, it is true that the complexity's internalization and externalization is dependent upon language use opportunities and time span, it is also used to differentiate first language (L1) users from second (L2) ones in English language education. The complexity's interpretation seems to have become pertinent in the context of too frequent contacts between L1 users and L2 users in English-speaking regions and between L2 users and L2 users for interactions, deliberations and discussions. Apart from its academic and scientific registering in today's use of English use in higher education, English use in academia has also been introduced to personalization, rationalization and linguistic sophistication. Nessler's (2020) consideration for the complexity's appreciation lies with linguistic maturity and sophistication. A simple production of a child reaches this maturity when an abstract thought of a higher degree of linguistic proficiency, of writing quality and of

—
grammatical development is produced for complex academic concepts in higher education (ibid.). Thus, it indicates that higher education tends to appreciate this higher degree of syntactic and linguistic complexity. Its representation of history in/though experience and event interpretation and of science through/in passivization are translated into noun phrasing and clause construction (Frantz et al, 2015).

READING

It is educationally true that a semantic development of this difference of the complexity is experienced and observed in the early years of human development. The sight paired with the bodies and the feelings and the senses of the language are further developed into a semiotic reasoning of language consistency. Obviously, reading habit that develops in childhood and passes its developmental phases in formative years of schooling is a life-long process. Pre-alphabetic experience of children tends to establish a relationship between the language's orthographic display of letters and phonological representations of (semantic) meaning in these years (Kirchmann *et al*, 2021). Their attempts to match phonemes with graphemes are recognized in the auditory working memory for reading fluency before they desire to read phrases and long threads of language (Adams, 1990, as cited in Kirchmann *et al*, 2021). However, it is also important to remember that sounds-spelling match is less likely to build successfully in some extraordinary linguistic conditions. Adults are engaged into reading practice at home and work too. Their professional need encourages them to read daily. The reader's eye movement has reportedly been observed to depict its relationship with language prosody and orthographic representation in phonological hub theory (Ashby, 2006, as cited in Ashby, 2006). Its stay at a word, which is called as the fixation, and its quick absorption of a good quantity of speech are used to indicate 'speech-like representation' of reading in a lips-sealed reading practice of an adult reader. This organization of speech absorption into verbal memory is initiated when the reader builds an orthographic image of a phonological representation of a word or a phrase at the fixation in the para(foveal) processing for working memory before s/he continues to interact with the text through para(foveal) process (*ibid.*) Pfof and Locher (2020) find a positive correlation between reading time and reading comprehension. Its fluency is found positively correlated with its comprehension in the literature of literacy (Kirchmann *et al*, 2021). The comprehension dependency lies with linguistic and cognitive processes of the reader (Elleman and Oslund, 2019). A skilled reader is observed with good working memory, comprehension monitoring, vocabulary, inferential ability, word reading ability and prior knowledge (*ibid.*). Manoli and Papadopoulou (2012) divide reading into strategies and skills. Reading strategies are employed to consciously plan to solve a reading problem and reach a good degree of comprehension while reading skills are automatic behaviour of the reader to effortlessly arrive at a text and its meaning. The

—

automation, which is usually facilitated with learning practice and repetition, is seen as a determiner of reading scholarship. Even though strategies and skills are used interchangeably, the former is reader-oriented mindful effort for a successful comprehension while the latter is a cognitive ability to interact with a text. Thus, reading skill is an ability to master if its strategy is adaptable to a response to a comprehension problem. Successful readers' metacognition is reportedly high among their peers in university education (Maasum and Maroof; 2012). Their awareness about effectiveness of reading strategies displays opportunities to repair and enhance comprehension. The reader with better metacognition is capable of deciding about reading focuses (global strategy), summarizing and circling reading information (support strategy) and slowing down the pace of reading and rereading (problem-solving strategy). Strategic knowledge or metacognition and comprehension's monitoring add to the skill. The reader's testing in modern testing centres is classed into degrees of reading proficiency (Kletzien, 1992) for its description in terms of text type today. A proficient reader uses his/her background reading and language knowledge to comprehend and differentiate texts of causation and comparison (*ibid.*) Grammatical information and previous reading knowledge seem to have added to the reader's comprehension of the subsequent reading text.

WRITING

The writing structure for a more informational flow for an argument formed among university practices of learning has a regular grammaticality. Academic writing has traditionally been stressed in the literature of language and education studies. Its reappearance in test designs has reportedly been assumed to be a social dimension to it in Bakhtin's framework. The framework's specification about the rejection of word use from a language dictionary for communication and the acceptance of the translation and conceptualization of past read or heard voices for an academic strengthening of a discourse lying across times seems to have introduced the significance of a discursal pattern (Constantinou *et al* 2018). The discursal plural centrality is further estimated in the test-taker's voice placed in a socially, culturally and institutionally controlled world of language. Thus, the test writer's language, which is authoritative and persuasive for community practice of communication, may find its way to the words of the test-takers. This academic grip of multivoiced discourse is frequently measured in student writing in higher educational institutions. Therefore, it is assumed that the epistemic stance of the writer demands a cohesive mental representation of both the writer and the reader (McNamara, 2013). The knowledge resulted with reading and writing is an easy qualification for coherence (*ibid.*). Self-efficacy, emotional intelligence and writing anxiety have also been used to

—

assess the grip over the subject in the language of graduate students (Huerta *et al*, 2017). The report that suggests that first language graduate students are far better at self-efficacy and emotional intelligence than second language ones is an indication of higher negative writing anxiety amongst non-English graduate students in Pakistan. Innovative assessment of short-written responses of non-English graduate students seems to have highlighted second language writing constraints in timed testing procedures in Hong Kong (Carles and Zhou, 2016).

THE VOICE

The easiest is the presence of the writer. It is a voice matching and differentiating at the level of some linguistic intelligence. AI is a more sum of such human linguistic intelligences conceptually matching. Authorial voice's promotion is subject to intellectual labour and academic liberty. Higher education in general and universities in particular have been considered as the places as institutions of academic integrity to help the students approach their scholarly drive in speech opportunities. To do so, they are encouraged to listen to and read experts and peers and cite references in discussions and writings. Despite academic writing's passivization in some scientific domain of discoveries, there is a growing trend of using personal pronouns to foster authorial voice in English writings. This is usually taken in school education before the students are guided to pursue their academic career in higher education. Carbone and Orellana (2010) have a report to share about a school's keenness to motivate its second language learners to style their English writings to address different readership on a scale of emotional distancing and academic structural consistency. Writing to a parent is reportedly intimate while to an expert it develops into structural persuasion for academic reasoning. The writer's attitudes and intellectual engagement are some of the developmental stages of persuasion and argumentation (Zhao and Wu, 2022). The student's originality of thought is usually foregrounded in English writing if he/she is aware of the connection between writer's writing and identity (Bird, 2013). A good writing curriculum of second language must have the learner's identity and its development into a human community for a communication tension and authorial claims for originality of thought, engagement and academic liberty (*ibid.*).

—

THE SCALE

How can one practise the developmental measure of the performance of each and every user of the language in the most consistent way for its historical richness? It is a rating scale or a rubric in the quickest and most continuous. The scale's use is not new to the process of English language test-takers' score's validation. International English Language Testing Service (IELTS), which is a high-stakes testing unit of the assessment of international students for English language proficiency, is frequently cited to comment upon the scale's descriptors and development in local contexts. The rubric's theoretical model of English language production of sentences in writing and utterances in speaking has been heavily dependent upon the thresholds set by Common European Framework of Reference for Languages (CEFR) of the Council of Europe. See CEFR for the development of communicative approach at Council of Europe (Council of Europe, 2022). Apart from its many types in other disciplines of studies, its two major types in English language assessment are analytic and holistic. Although its analytic type, which has separate identifiable assessing objectives for different sections of language description and proficiency, is popular among IELTS raters or testers, its closely holistic, which cannot be considered as a typical impressionistic, has also been tested with Aptish (English tests for organizations) in Pakistan (British Council, 2022; Speaking and Writing Rating Scales Revision, 2017). Lin and Liu's (2022) introduction of an analytic scale to assess basic paraphrasing skill of English language users of a university in China is an instance of organizational need of language achievement, measurement and problem diagnosis in English language description in non-English speaking countries. It is observed that language scale developers tend to maintain good relationships of pragmatic choice of words, syntactic development of a thought, phonological exhibition of speech and changing patterns of lexical and phrasal meanings in contextual dependent clusters of interactants. Charles Alderson and Tim McNamara for measurement of language assessment are crucial. The judgement or effect of the rater, test item difficulty and test-taker's performance have been found variable at the scale's development and operation. Rasch measurement is increasingly used to highlight this complexity of language, its users and its testers in recent literature of language studies (Aryadoust and Sayama, 2021).

THE PERSPECTIVE

Reacting to a test item or a subjective question in a timed written examination requires a test-taker to demonstrate their ability to see and comprehend. In order to incline to comprehension, they are expected to have attained required proficiency in

—

reading and its various functions within language. Although reading is sight-assisted, comprehension is run through mental capacity to gauge degree/level of task difficulty and translation of knowledge tailored to meet expectations of expert readers or examiners. For the sake of ease and knowledge economy, I have utilized the word 'understanding' to witness it, as comprehension results into various levels of understanding. By 'expectations' I mean to say if the examiner as a reader is eager to learn and discover the examinee's intellectual bent towards the subject or is merely anxious about translation of information retrieved from classroom lectures and textbooks into the given/required language. By a proficient reader I mean to say that the examinee's ability of sight has been trained to differentiate words and read. They must be capable of differentiating description from evaluation, discussion from definition, and commentary from illustration. The expectations may vary among the examiners in the absence of a rubric or a rating scale. The examinee who can read and elicit information from the words utilized to support the description of the task or the test item is expected to be a producer of expressions of basic understanding. If they have a better understanding of the language and its user, a better written response may be anticipated. If a classroom teacher's role has changed to the examiner, or a group of subject experts has been asked to write tasks, frame questions, set difficulties and evaluate, or a panel of testing within the domain of the subject or the specialization has been tasked to take care of both academic and administrative administrations of the test process as tasks or as written tests, a positive backwash effect may be observed.

Besides motivations to a linear production of sentences punctuated by paragraphs, the examinee's intellect flourishes in a state of academic reasoning. Its progression reaches in a setting of (tertiary) education. Since an argument is a piece of writing aimed to determine validity and reliability of core textbooks' theories, issues and perspectives in archived form to take information into knowledge composition, the examinee as a purpose-specific and socially placed writer is responsible for differentiation of facts and opinions, exemplification of events and issues and description before he/she writes an account of critical appreciation. As a result, the argument is socially situated and academically written. By socially situated I mean to say a theory or an issue is coherent only if its interpretation is executed within systems of language and the users of speech communities. By academically written I want to refer to the structure and the language of the critique. Apart from a critical commentary and a conclusion, the argument is augmented with rich expert vocabulary, an authorial voice and a noticeable continuation of one or many methods of assessment.

Language and thought are two prerequisites of describability. No notion, issue or phenomenon is describable without (a) language in a beyond art formulation for a

—

more behavioural attitudinalization and human fitness for speech communities. By descriptability I mean to refer to the capacities of language and their utilization by its users from a meaningful display of thought. There is an agreement among a group of linguists and language philosophers that concept learning is either synonymous with language learning or inseparable from it. Language learners are exposed to physical settings and themes ranging from the universe to life cycle. If it is assumed that the universe is an encapsulation of language, language users are called successful artists, scientists, thinkers, theologians and writers. Their expressions contain themes of time and space, life and death, love and hate and theories of modernism to minimalism, existentialism to spiritualism and geometry to geology. Each theory or theme can be described once it is supplied with huge sets of vocabulary items produced by/through human beings and their regular interactions with subject matters. The items are compiled by lexicographers, promoted by grammarians and utilized by the users. The learners are encouraged to take advantage of recursive mode of language and use its specialist vocabulary of business, law, science, creativity, education and religion in respective settings. Sets of expert vocabulary items are deliberated and designed for fields of knowledge and differentiated from specialist vocabulary. If a carpenter is in possession of specialist vocabulary items of carpentry, an engineer is expected to be capable of using both types.

Analytical capacity of a human mind can be translated into/through writing. The human brain maps up thinking patterns subsequent to the development of writer's analytical capability. His/her aptitude is stimulated with a supply of adequate information coupled with illustrations. The reception that is usually carried out by/through the senses of sight and hearing is processed in the brain and transmitted through the upper limbs into carefully selected words. The process is an examination of the information that is further developed with the writer's act of reflection comparing past events with present moments of reality for a course of action in the future. The careful selection is, in fact, a manifestation of the same cognitive process of rationalization achieved through elicitation, exemplification, description and interpretation. The writer as an interpreter reflects and refers back to documented evidences of background/tacit knowledge. A thought's or an issue's simplification is his/her intellect's utilization to produce a measurable result and conclusion. Successful operation of this theoretical mechanism of the researcher may result into a merger of information into the writer's tacit knowledge about the subject matter displayed in written paragraphs.

As it is just reported, the authorial voice is an improving sign contributing to the recognition of the writer's intellectual rigour. I do not think readers are less interested to know about explicit and implicit markers of the writer's intellectual stances in the age of democratic values and international citizenship. Probably this is why both

—

active and passive participations of the writer are regularly observed in well-written pieces of writing in international systems of writing tests. The continuation of participation ranges from passivity (that is, 'I learn') to activity (that is, 'I want'). At times, they are encouraged to support or oppose an issue of general concern and strengthen their opinion with an array of facts. It is true that they open an account of an unbiased description but need to draw a conclusion using a method of evaluation that is comprised of opinions and self-checklists. The use of first-person singular along with a verb is an indicator of writer's feeling, intensity of participation, origin and identity group.

Grammatically correct sentences are a representative mode of sense making. The choice of correct words and the use of correct structures are crucial to signify the quality of message and its interpretation. As the language of English has a huge lexicon its writers are expected to demonstrate his/her skill to choose words suitable to their collocative environment to determine differentiation and similarities using words of their respective semantic fields and affinities, and to word their expressions compatible with the style of written speech. As writing is a combination of structures, dangling modifiers, incorrect tenses, redundancies, misspellings and punctuation mistakes are learned as typical instances of grammatical inadequacies. Variation in English sense-making process is detectable from one region to another and from native-English speaker to nonnative English speaker. By 'grammatically correct' I would like to refer to teaching models of (standard) varieties of English used for communication in the western hemisphere. However, I do understand social and sociolinguistic values attached to the regional varieties of the language and their significance in enriching English pragmatic function or making it dysfunctional within and across the region(s). Since the users of English in Pakistan have been exposed to Standard British English and Standard American English in institutions of formal education, they are expected to express their thoughts and feelings in the second language or the language of their preference.

A piece of writing is of considerable worth if it is supplemented with examples and citations. An experienced writer does not hesitate to participate as an exemplary figure if he/she thinks that such inclusions of examples may help personalize and verbalise the issue or the thought in question. His/her writings contain personal expressions of popular personalities' and theirs. They are helpful for a connection between the reader and the writer. The piece of writing is further supported or strengthened or displayed with academic integrity if it is cited where necessary. Not only does it help retain academic integrity but it also helps the writers separate his/her opinions and examples from facts. The writer has skills to incorporate such examples and references and monitor their influence over the description and the evaluation of the written response. The influence should be positive, not overwhelming. By

—
overwhelming, I mean to say that their contributions must not push other constituents of knowledge-driven information to follow. On the contrary, their application in the writing would be a mere follow-up or an add-on.

A great amount of work has been done on the validity of language use and its testing. One should read SAGE and some good virtual places of the research. We have just seen the nearest British Council and the Educational Testing Service (ETS). Validity of facts is a challenging job. It is rather difficult to estimate if a piece of information is fabricated, concocted, false or deceptive, as every second user of English is aware of an abundance of information supplied to websites daily. Thousands of them are either modified or replaced on hour basis. Millions of cheap quality books are published and shipped. Disorientation among language learners as concept learners needs to be realized with an assessment of facts. The analytical value of the description of the issue needs to be matched or compared with the source of information. Both ways of content generation may be viewed against the writer's personal experiences and observations. Classroom reading texts, recommended books and discussions are a help for required validity.

A Successful application of the descriptors for extended written responses is foiled by inadequacies in an area or/and phase of a paper-pencil examination process. They may be characterized by unsupportive physical setting, untrained administrative staff and test-taker's dissatisfaction over their own written attempts under stress. Pakistan is a place where inadequacies haunt professionals striving to standardize local educational practices and motivate learners as scholars for excellence. In fact, an American graduate, Michelle Kemp, is unsure about the purpose of academic tests under stress:

Tests show what you can do under stress, but during the latter half of the exam, you are so exhausted from three hours of testing that you can really hurt your scores (Kempt; 1982:113)

Her complaint has been resolved in the ETS (for further reading, see Raza, 2007). Similarly, most educational institutions in and outside Pakistan have introduced writing assignments for enhanced scholarship of classroom practices. Tariq Rahman's (1999) and Sabiha Mansoor's (2005) writings about prospects and problems of English language education in Pakistan are rare reads that help understand contexts and contexts of knowledge promoted through curricula and classrooms. Besides the description, the writer's response's effectiveness needs to be re-evaluated. How sincere and determined is the writer must not go unnoticed. A paragraph in plain language is far more authentic than a lengthy piece of flowery writing.

THE STUDY

The study was quasi-research. The researcher was requested to facilitate a group of twenty students to study World Englishes in a four-month course of a two-year master's programme in a public institution of higher education of Pakistan in 2018. In order to improve the performance and see a need to bring about educational reforms, a plan was chartered out to begin with the development of a nine-descriptor scale of rating and its theoretical justification. It was more action research in a classroom practice-initiated work.

Action research is a process of investigating one's practice to understand and improve it. Kemmis, McTaggart, and Ratallick (2004) define action research as 'a form of collective self-reflective enquiry undertaken by participants in social situation in order to improve the rationality and justice of their own social or educational practices, as well as their understanding of these practices and the situations in which these practices are carried out. (Dean; 2009:87)

The knowledge structure is so immersive in such conditions of learning that no teacher as an administrator of a part of it can leave a growing mass of enquiry in the city.

Teachers' actions ... (may require) ... further improvement, and improvement requires involvement in the process of inquiry. Action research therefore aims to investigate reality in order to change it and at the same time, it aims to change the reality in order to investigate it. (Mohammad and Kumari; 2004:112)

The group was instructed to submit a written assignment of an essay (almost two months were given for the completion of the seventh task) and write extended written responses to a set of six tasks/questions (See HANDWRITTEN TEST PAPER) in a three-hour paper-pencil test in a teaching room of the institution. Task's doability's duration was checked. The responses were read and rated.

-

HANDWRITTEN TEST PAPER

ADDENDUM 01: TEST ITEMS IN THREE HOURS

Subject: World Englishes

Code: 652-13

Duration: Three hours (04PM-07PM)

Marks: One hundred

Programme: M.A. English (Ling.)

Department: English

Faculty: Wayne Raza

Date: 20 December 2018

Attempt five test items. Each item carries equal marks.

1. Write a description ^{of} Kachru's interpretations of the concentric circles.
2. Why is there a growing consensus among a group of linguists that the usage of English as an international ^{language} will be replaced with a gradual fall with the passage of time? Write reasons and illustrate them with global trends of translation, technology and community identities as ^{few} contributing factors or indicators of the language's replacement.
3. Write a note on the theoretical framework(s) of language proficiency in establishing or refuting the subject of World Englishes.
refuting

1 →

4. How far is Jenkins' model of phonology of English as an international language a realistic approach towards understanding language? Discuss.
5. Establish an argument in support or opposition of the subject as a (philosophical) way to appreciate language change and its consequences. Write reasons and examples.
6. Write short notes on any two of the following terms:
 - (i) eurocentricity
 - (ii) modernism
 - (iii) linguistic deficit
 - (iv) nativeness

WHAT THE RATER LOOKS INTO

Out of the nine descriptors of the scale on numerical values one to ten, the set of data in the form of scores ranging between the values was helpful for a suggestion to predicting difficulty levels of the test-items and the descriptors. It is assumed that the test taker's performance was positively related with the level of difficulty of the items and the descriptors. If his/her score was excellent in a test item, the item was the least difficult. The fourth and the sixth descriptors of the scale were the least difficult to score while the fourth and the sixth were relatively more difficult. The remaining descriptors were almost equally moderate in terms of level of difficulty. The concentration of the values of 10 in the first descriptor and of 07 and 08 in the sixth descriptor is a display of excellent demonstration of the test-taker's proficiency in writing. See SCORES, DESCRIPTORS AND TEST ITEMS.

SCORES, DESCRIPTORS AND TEST ITEMS



1 : SCORES, DESCRIPTORS AND TEST ITEMS

D5										D6										D7									
06	05	04	07	06	05	08	06	07	04	05	06	08	06	08	08	08	08	04	07	06	05	06	04	06	06	05	07	07	06
08	04	06	08	05	08	05	08	06	06	06	06	08	08	08	08	08	08	07	07	05	06	07	05	06	05	07	07	04	06
07	08	06	06	08	07	07	05	06	07	07	07	08	08	08	08	08	08	06	04	05	06	05	06	06	04	05	07	07	06
05	08	04	08	08	07	07	07	08	07	08	07	09	09	09	09	09	09	07	08	04	04	05	08	08	05	05	05	05	05
05	06	08	07	08	06	05	04	01	02	06	09	08	08	08	08	08	08	06	07	04	05	06	06	05	06	04	04	05	07
06	05	07	03	04	06	09	04	07	07	07	07	08	08	08	08	08	08	06	07	05	04	04	05	06	08	06	07	07	06

1515088 1550572

1 : SCORES, DESCRIPTORS AND TEST ITEMS

D5										D6										D7									
06	05	04	07	06	05	08	06	07	04	05	06	08	06	08	08	08	08	04	07	06	05	06	04	06	06	05	07	07	06
08	04	06	08	05	08	05	08	06	06	06	06	08	08	08	08	08	08	07	07	05	06	07	05	06	05	07	07	04	06
07	08	06	06	08	07	07	05	06	07	07	07	08	08	08	08	08	08	06	04	05	06	05	06	06	04	05	07	07	06
05	08	04	08	08	07	07	07	08	07	08	07	09	09	09	09	09	09	07	08	04	04	05	08	08	05	05	05	05	05
05	06	08	07	08	06	05	04	01	02	06	09	08	08	08	08	08	08	06	07	04	05	06	06	05	06	04	04	05	07
06	05	07	03	04	06	09	04	07	07	07	07	08	08	08	08	08	08	06	07	05	04	04	05	06	08	06	07	07	06

1515088 1550572

T refers to test item.
D refers to descriptor.
The circles ones are misfits.

HOW THE SCALE WORKED

1515002

The last item has been introduced and concluded appropriately. The scheme of argumentation is a composition of an assessment of the issues, theories, views or/and terms' value.

3. The description is appropriately worded and academically strengthened with textual evidences. The response is a fair representation of necessary information.

4. Background knowledge, systematic arrangement of explicit opinion(s), rationalisation, reflection(s), and intellectual rigour are main strengths of the analysis.

5. The writer has made several attempts to establish their intellectual stance and maintain it throughout the response.

6. In addition to usage of right words, writing style adopted/used informed by the writer is an excellent reinforcement of over-free grammar.

7. The instances are comprised of excellent personal examples, relevant experiences and observations and reliable bibliographic references.

8. The facts are valid.

9. The quality of the response is impressive.

	01	02	03	04	05	06	07	08
10				07	10	08		10
08				07	04	05		06
08				07	04	05		08
05				06	04	06		05
07				06	04	07		05
08				07	07	08		08
06				06	04	06		05
07				07	05	07		06
07				06	04	04		07

The scale used to add numerical value to the respondent 1515002

However, expressions for analysis and discipline-based knowledge were meager among them, as almost all test-takers' score were reported under the ranges of the last two classes of the scores for the fourth and seventh descriptors. Probably a recollection of thinking and studying are hard to find places in the responses of this timed-achievement test in a physical setting of a room in a state of dilapidation. Despite this oddity, the test-takers' performance was surprisingly good. Their scores in the second, third and fifth descriptors were a fair representation of authorial voice and content knowledge. See HOW THE SCALE WORKED. Occurrences of the scores 05 to 08 for the last two descriptors were an implicit gesture of academic honesty and integrity. Except the sixth, all tasks were reported with moderate level of difficulty.

—

There was no difference of difficulty found between the seventh and the set of timed-tasks under invigilation. The sixth task was left incomplete or attempted incorrectly.

There are four instances of extraordinary scores. 1515018 scoring three in the fifth descriptor of the test item containing the working paper had an unnoticeable writer's opinion. See HOW THE SCALE WORKED for the respondent/test-taker 1515002 again. On the contrary, the examinee 1515058 scoring nine in the same descriptor of the same item retained an unbelievable visibility of authorial voice and intellectual stance. The third instance is about excellent demonstration of use of English grammar by 1515018 in the fifth test item projecting an angle of argumentation. The response of 1515048 to short notes was against the condition of academic honesty stated in the ninth descriptor.

SUGGESTIONS

Firstly, if I go back to the issue of written responses' length in timed tests, I want educationists and testers to revisit educational policies and procedures and learning assessment. The absence of the support for the testing and its fairness is likely to increase a distrust among the community of the scholars and students. Secondly, if academic tests' language is not a variable in higher education (such a postulation is beyond reasoning under prevailing environment of the country), tasks' or test-items' effectiveness can be accurately measured with rating scales, scores and users. Thirdly, the action research was a tool to estimate a shift or a positive change within individual capacity. The scale's application to social sciences' (extended) written responses may help educationists incentivize academic administration of programmes of higher education. Fourthly, the scale's accuracy and validity need to be measured by a group of raters on several occasions of testing environment over a period of five years. Lastly, the participants should have been requested to show their consent before data collection.

BIBLIOGRAPHY

Adams, MJ (1990). *Beginning to read: Thinking and learning about print*. The MIT Press: US.

Alderson, J C (2007). The challenge of diagnostic testing: do we know what we are measuring. In J. D. Fox (Ed.), *Language testing reconsidered* (pp. 21-39). University of Ottawa Press: US.

Alderson, C J (2003). *Assessing reading*. Cambridge University Press: UK.

Aryadoust, V, Ng, L Y, & Sayama, H (2021). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing*, 38(1), 6–40. DOI: <https://doi.org/10.1177/0265532220927487>

Ashby, J (2006), Prosody in skilled silent reading: evidence from eye movements. *Journal of Research in Reading*, 29: 318-333. DOI: <https://doi.org/10.1111/j.1467-9817.2006.00311.x>

Ashby, J (2006). The nature of phonological representation in reading: Evidence from eye movements and event-related potentials. *PhD Thesis*. University of Massachusetts: US.

Bachman, L F (1991). What Does Language Testing Have to Offer? *TESOL Quarterly*, Winter, 25(4), 671-704. DOI: <https://doi.org/10.2307/3587082>

Bird, B (2013). A basic writing course design to promote writer identity: three analyses of student paper. *Journal of Basic Writing* 32 (1), 62-96.

—

Brunfaut, T, Kormos, J, Michel, M, & Ratajczak, M (2021). Testing young foreign language learners' reading comprehension: Exploring the effects of working memory, grade level, and reading task. *Language Testing*, 38(3), 356-377.
DOI: <https://doi.org/10.1177/0265532221991480> (Original work published 2021)

British Council (2023). British Council. <https://www.britishcouncil.org/>

Buck, G (2003). Assessing listening. Cambridge University Press: UK.

Carbone, Pand h Orellana, M (2010). Developing academic identities: persuasive writing as a tool to strengthen emergent academic identities. *Research in the Teaching of English* 44 (3), 292-316.

Carless, D & Zhou, J (2016) Starting small in assessment change: short in-class written responses. *Assessment & Evaluation in Higher Education*, 41 (7), 1114-1127.
DOI: 10.1080/02602938.2015.1068272

Chiony, S (2012). Saving face. Retrieved from <https://www.imdb.com/title/tt2140371/>

Constantinou, F, Crisp, V, & Johnson, M (2017). Multiple voices in tests: towards a macro theory of test writing. *Cambridge Journal of Education*, 48(4), 411-426.
<https://doi.org/10.1080/0305764X.2017.1337723>

Council of Europe. (2022). Council of Europe. Retrieved 7 December 2022 from <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>

Dimova, S, Yan, X, & Ginther, A (2022). Local tests, local contexts. *Language Testing*, 39(3), 341-354. DOI: <https://doi.org/10.1177/0265532221092392> (Original work published 2022,

Durbeen (2024). Durbeen. Retrieved from durbeen.org

Eastwood, J (1999). Oxford Practice Grammar. OUP: Pakistan.

Elleman, A M, & Oslund, E L (2019). Reading Comprehension Research: Implications for Practice and Policy. *Policy Insights from the Behavioral and Brain Sciences*, 6(1), 3-11.
DOI: <https://doi.org/10.1177/2372732218816339>

ETS (2023). English Testing Service. <https://www.ets.org/>

Foreign Service Institute (2023). Key topics. <https://www.state.gov/key-topics-foreign-service-institute/>

–

Frantz, R S, Starr, L E, & Bailey, A L (2015). Syntactic Complexity as an Aspect of Text Complexity. *Educational Researcher*, 44(7), 387-393.

DOI: <https://doi.org/10.3102/0013189X15603980>

Grande, B, Leishman, C D, Skilleås, H K (2022). "I can't come to the words": Effects of including military flavour when testing the oral proficiency of Norwegian joint terminal attack controllers. *English for Specific Purposes*, 68, 73-86.

Heaton, J B (1979). *Writing English language Tests*. Longman Handbooks for Language Teachers. UK.

Huerta, M, Goodson, P, Beigi, M & Chlup, D (2017). Graduate students as academic writers: writing anxiety, self-efficacy and emotional intelligence. *Higher Education Research & Development*, 36 (4), 716-729. DOI: 10.1080/07294360.2016.1238881

Hungerland, I C (1955). The Interpretation of Poetry. *The Journal of Aesthetics and Art Criticism*, 13(3), 351–359. <https://doi.org/10.2307/426435>

Jin, Y (2022). Test-taker insights for language assessment policies and practices. *Language Testing*, 40(1), 193-203. DOI:

<https://doi.org/10.1177/02655322221117136> (Original work published 2023)

Khattak, SG (2012). Assessment in schools in Pakistan. UNSECO. Retrieved from https://neqmap.bangkok.unesco.org/wp-content/uploads/2019/08/6_Assessment-in-schools-in-Pakistan.pdf

Košak-Babuder, M, Kormos, J, Ratajczak, M, & Pižorn, K (2018). The effect of read-aloud assistance on the text comprehension of dyslexic and non-dyslexic English language learners. *Language Testing*, 36(1), 51-75.

DOI: <https://doi.org/10.1177/0265532218756946> (Original work published 2019)

Kletzien, S B (1992). Proficient and Less Proficient Comprehenders' Strategy use for Different Top-Level Structures. *Journal of Reading Behavior*, 24(2), 191-215. DOI:

<https://doi.org/10.1080/10862969209547772>

Kirschmann, N, Lenhard, W, and Suggate, S (2021). Influences from working memory, word and sentence reading on passage comprehension and teacher ratings. *Journal of Research in Reading*, 44: 817–836. DOI: <https://doi.org/10.1111/1467-9817.12373>.

Jin, Y (2022). Test-taker insights for language assessment policies and practices. *Language Testing*, 40(1), 193-203.

DOI: <https://doi.org/10.1177/02655322221117136> (Original work published 2023)

—
Liu, Sen and Lin, Dunlai (2022). Developing and validating an analytic rating scale for a paraphrase task. *Assessing Writing* 53 (2022). 1-17.
DOI: <https://doi.org/10.1016/j.asw.2022.100646>

Locher, F, and Pfost, M (2020). The relation between time spent reading and reading comprehension throughout the life course. *Journal of Research in Reading*, 43: 57–77.
DOI: <https://doi.org/10.1111/1467-9817.12289>.

Luoma, S (2008). *Assessing speaking*. Cambridge University Press: UK.

Polyxeni, M& Papadopoulou, M (2012). Reading Strategies Versus Reading Skills: Two Faces of the Same Coin. *Procedia - Social and Behavioral Sciences*. 46, 817-821.
DOI: <https://doi.org/10.1016/j.sbspro.2012.05.205>.

Maasum, T N R, Mohamad, T&Nooreiny, M (2012). Empowering ESL Readers with Metacognitive Reading Strategies. *Procedia - Social and Behavioral Sciences*. 69. 1250-1258. DOI: <https://doi.org/10.1016/j.sbspro.2012.12.058>.

Marklund, U, Lacerda, F, Persson, A&Lohmander, A (2018). The development of a vocabulary for PEEPS-SE—profiles of early expressive phonological skills for Swedish. *Clinical Linguistics & Phonetics*, 32(9), 844-859.
DOI: 10.1080/02699206.2018.1445780

Martinet, A (2006). Functional Grammar: Encyclopedia of Language & Linguistics. *Elsevier*, 677-682.
DOI: <https://doi.org/10.1016/B0-08-044854-2/02036-8>.

McNamara, D S (2013). The epistemic stance between the author and reader: A driving force in the cohesion of text and writing. *Discourse Studies*, 15(5), 579-595.
DOI: <https://doi.org/10.1177/1461445613501446>

Nasseri, M (2021). Is postgraduate English academic writing more clausal or phrasal? Syntactic complexification at the crossroads of genre, proficiency, and statistical modelling. *Journal of English for Academic Purposes*, 49.
DOI: <https://doi.org/10.1016/j.jeap.2020.100940>.

North, B (2007). The CEFR Illustrative Descriptor Scales. *The Modern Language Journal*, 91(4), 656–659. DOI: <http://www.jstor.org/stable/4626092>

Ramanathan, H (2008). Testing of English in India: A developing concept. *Language Testing*, 25(1), 111-126. DOI: <https://doi.org/10.1177/0265532207083747> (Original work published 2008)

Rea-Dickins, P (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249-258. DOI: <https://doi.org/10.1191/0265532204lt283ed> (Original work published 2004)

Red (2024). www.red.org

Speaking and Writing Rating Scales Revision (2017). Technical Report: Speaking and Writing Rating Scales Revision TR/2017/001. English language assessment research group.

British Council. Retrieved 7 December 2022 from https://www.britishcouncil.org/sites/default/files/aptis_scale_revision_layout.pdf

Tarai zameen par (2007). Retrieved from <https://www.imdb.com/title/tt0986264/>

The British Broadcasting Corporation (2024). Skimflation. Retrieved from <https://www.bbc.com/worklife/article/20231003-skimflation-an-even-sneakier-form-of-shrinkflation>

The language of literature (2005). McGougal Little. Houghton Mifflin Company: US.

University of Waterloo (2024). Teaching and testing. Retrieved from <https://uwaterloo.ca/centre-for-teaching-excellence/catalogs/tip-sheets/designing-multiple-choice-questions#:~:text=A%20multiple-choice%20question%20%28MCQ,incorrect%20answers%20to%20the%20question.>

Weigle, S C (2002). *Assessing Writing*. Cambridge University Press: Cambridge.

Wright, J (2003). *Dictionaries*. Oxford University Press: UK.

Zhao, C and Wu, J (2022). Perceptions of authorial voice: Why discrepancies exist. *Assessing Writing*, 53, 1-12.

Zhao, H., & Zhao, B. (2020). Co-constructing the assessment criteria for EFL writing by instructors and students: A participative approach to constructively aligning the CEFR, curricula, teaching and learning. *Language Teaching Research*, 27(3), 765-793. DOI: <https://doi.org/10.1177/1362168820948458> (Original work published 2023)

-

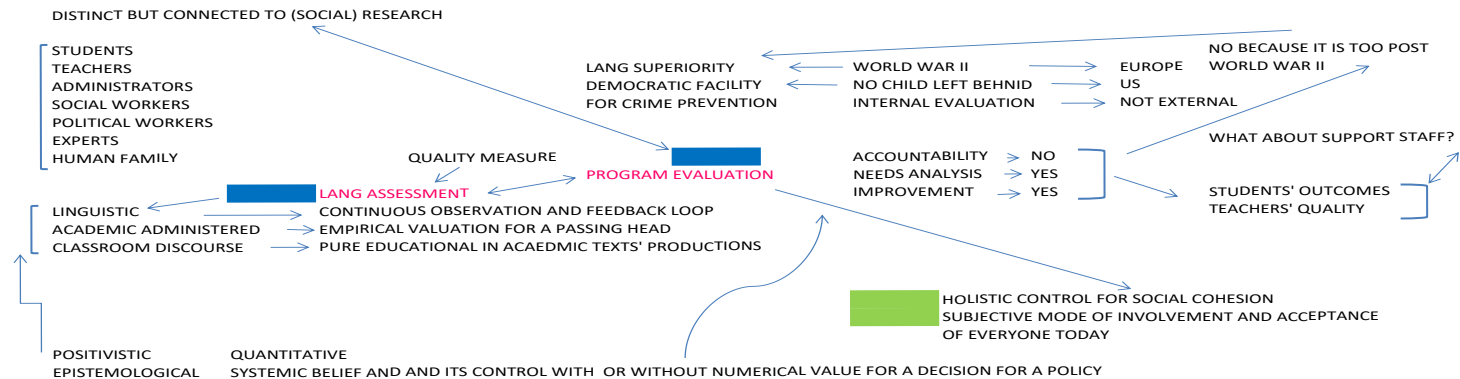
NOTE-TAKING

TALKS 1, 14, 15, AND 16

COURSE LANG ASSESSMENT AND PROGRAMME EVALUATION
TALK 1
READ NORRIS' LANGUAGE PROGRAM EVALUATION
GEORGETOWN UNIVERSITY
2016
TF WAJDAN R

LISTEN TO THE TALK AND CONNECT.

TALK 1



LECTURE
WORKSHEET
TEXTS
REFERENCE

VOCABULARY ITEMS
FILL-IN AND CONNECT
LINGUISTICS
HEATON'S
MARKLUND'S

TALK 14

WORD
WEB LINK
AUDIO CLIP

RED
<https://www.red.org/>
WR

for representation and prevention

EARLY YEARS

AERODYNAMIC SPELLING SYLLABIFICATION

CONSONANTALS

PHONO MORPHO SYNTACTIC GRAPHEMIC

PARENTAL INTERVENTIONS

VOCABULARY

CONNOTAT

LEXICAL DEVELOPMENT

SEMANTIC DEVELOPMENT

REFERENTIAL
SOCIAL
AFFECTIVE
METAPHORICAL

DENOTATIV

ADULTS

ACTIVE/PASSIVE

SPECIALIZATION

DAILY LIFE

FOREIGN

PAKISTAN

INDIA

to be like a red to a bull
be in the red

LOCAL

US

SWEDEN

LECTURE	LANG TEST'S OPERATION	
WORKSHEET	FILL-IN AND CONNECT	
TEXTS	LINGUISTICS	
REFERENCE	HEATON'S	WITH M M PHOTOCOPIER
	BACHMAN'S	WITH THE GOOGLE CLASSROOM
	RIGHT'S	PHOTOCOPIABLE HERE
WORD	KNOWLEDGE BASED ECONOMY	
WEB LINK	https://www.adb.org/sites/default/files/public for representation and progress	
AUDIO CLIP	WR	

TALK 15

CLASSROOM TALK 15

Bachman's theoretical complexity and test-taker's characteristics				Bachman's statistical measure				Bachman's communicative approach							
LANGUAGE TEST'S OPERATION												Bachman's material sense with Right's Dictionaries			
validity				reliability				intrarater interrater							
test-taker's abilities' dependent interactional				test score evenness											
authenticity				construct validity teaching-testing											
situational															
tester's and administrator's abilities' dependent															
				language use-language test											
empirical validity															
concurrent								theoretical display							
predictive															
language skills				imitation				learning							
knowledge skills				processing				creating							
				industrializing											
FUTURE PAKISTAN															
PAKISTAN ZINDABAD															

FUTURE PAKISTAN
PAKISTAN ZINDABAD

COURSE LANGUAGE ASSESSMENT AND PROGRAMME EVALUATION
DEPARTEMENT ENGLISH
TEACHING DEPARTMENT ENGLISH
PROGRAMME BS/MA
TEACHING FACULTY WR
TOPIC PROGRAMME'S EVALUATION'S SECTIONAL DESCRIPTION
READ HEATON'S

TALK 16

CLASSROOM TALK 16

There are at least three simple ways to the improvement of the test procedure's evaluation.

- 1 Measure the test item difficulty.
- 2 Measure the rater's severity.
- 3 Measure the test-taker.

That is, how numerically easy/difficult it is to the test-taker
That is, how numerically severed/simple the rater/reader is to the test-taker's test item.
That is, how numerically a good/bad the test-taker is to the test item.

This is how the measurement can be initiated:

- a Collect the numerical values.
- b Do some descriptive statistical analysis.
- c Do some action research plan.
- d Have it funded.
- e Publish and present.
- f Improve or continue.



frequency development TRENDING MARKER range SIGNIFICANT RANGE central tendency HIGHER MEAN standard deviation SPREAD REDUCES NEGATIVE BIAS.



READ Heaton's Writing English Language Tests' Chapter 11.